科学佳作 科学力作 科学妙作

支点丛书

[英] 约翰·黑格/著
by John Haigh
李大强/译

吉林人民出版社



机会的数学原理

Taking Chances

明知其输而博赢的概率分析



[英] 约翰·黑格 著　李大强 译

机会的数学原理



吉林人民出版社

Taking chances

by John haigh

本书由 Oxford University Press 1999 年版翻译

吉林省版权局著作权合同登记

图字:07—2000—382

(吉)新登字 01 号

机会的数学原理——明知其输而博赢的概率分析

著　　者　[英]约翰·黑格

译　　者　李大强

责任编辑　范春萍　黄式刚　　　封面设计　翁立涛

责任校对　黄式刚　　　　　　　版式设计　胡学军

出 版 者　吉林人民出版社　0431—5649710

　　　　　(长春市人民大街 124 号　邮编 130021)

发 行 者　吉林人民出版社

制 版 者　吉林人民出版社激光照排中心　0431—5637018

印 刷 者　长春市人民印刷材料厂

开　　本　850×1168　1/32

印　　张　14

字　　数　300 千字

版　　次　2001 年 8 月第 1 版

印　　次　2001 年 8 月第 1 次印刷

印　　数　1—6 200 册

标准书号　ISBN 7-206-03818-2/F·791

定　　价　26.00 元

如图书有印装质量问题,请与承印工厂联系。

内容简介

概率与我们的日常生活息息相关。当我们过马路的时候，当我们上保险的时候，当我们买彩票的时候，我们都在和不确定性打交道。然而，普通人对概率所知甚少。在我们关于概率的知识中，有许多本应避免的错误。本书的目的是让普通人获得应用概率知识的能力。书中深入探讨了彩票、轮盘赌、扑克游戏等以概率为核心的问题，引人入胜的分析经常使读者茅塞顿开。你可以把它当做一本精妙的小说，也可以把它当做一本实战指导手册。

前　言

几乎每一个你有意识地作出的决定都与概率相关。当你穿衣服时，你的决定取决于你对天气的判断；当你过马路时，你的决定取决于你对发生车祸的可能性的估计，你储备备用灯泡，是为了应付某种可能性；你向保险公司投保，理由是“以防万一”。对于概率，人类一定拥有非常充分的直觉，否则人类文明不可能演化到现在的状态。

另一方面，尽管人类对于概率有非常好的直觉，多数人对概率的理解是不充分的。最显著的问题有两个：其一，某些事件发生的概率值极低，如何定量地分析这些很小的概率值之间的差别？比较两个判断。A：“事件 a 发生的概率为 1/1 000”；B：“事件 a 发生的概率为 1/1 000 000”。通常人们以为这两句话的意义是一样的，都是说事件 a 发生的可能性很小。事实上，这两个概率值不同，前者是后者的 1 000 倍；其二，面对具体的判断时，如何避免错误信息的干扰？一个小例子：我给你一张美女照片，你的任务是猜测此人的职业：模特还是职员？很多人会猜前者。实际上，模特的数量比职员的数量少得多，所以，从概率上说这种判断是不明智的。

当你买彩票时，当你玩扑克游戏或骰子游戏时，当你赌马时，你清醒地意识到你在和随机事件打交道。即使你不懂如何计算概率，经验和直觉也能帮助你作出判断。在很多场合，这些粗糙的判断是简单有效的；也有很多场合，除非经过缜密的分析和精确的计算，你的结论会错得离谱。假设你作为陪审团的一名成员出庭，被告被指控犯有谋杀（或绑架）罪。法庭掌握了一些事实。如果被告是无辜的，则这些事实发生的概率为100万分之一。有些人可能认为这个判断等同于“在这些事实已发生的前提下，‘被告是无辜的’的概率为100万分之一”。大错特错！请比较以下两个判断：“珍妮在下雨天遗忘外套的概率为1%”以及“在珍妮遗忘外套的时候正在下雨的概率为1%”。二者的含义显然不同。

vii

在本书中，我介绍了许多与机会有关的问题，我的着眼点是趣味性、启发性和知识性。有些例子读者很熟悉，有些例子读者会觉得陌生，全部内容都与概率相关。每一章围绕统一的主题，某些章后面有附言，附言是对某些概率论方法的发挥。通常，我们的分析方法可以适用于两（或多）个人（或队伍）对抗的游戏。如果你认真地分析其中的概率问题，你往往会在游戏中获利——通常这就是吸引我们关注这些问题的原因。

本书是为初学者写的，希望没受过严格的科学和数学训练的读者可以接受它。据说，在通俗读物中每增加一条数学公式就会赶走一半读者，所以我在正文中尽量避免符号化的表达方式。当然，我不能删除所有的公式。书末有5个附录，其中涉及到比较复杂的数学知识。我在正文中加了一些方框，方框中的文字通常是比较复杂的数学分析。对数学没兴趣的读者可以略过这些内容。我希望读者不会因此感到不便。如果你认真地读下去，你会发现很多精妙的思想其实惊人地简单。

为便于阅读，我对许多数字取了整。当我说某个概率值是50%，这个值有可能是精确值（比如抛一枚硬币，正面向上的概率是50%），也有可能是近似值（比如在某场球赛中主队获胜的概率为50%）。通过上下文可以区分这两种情况。通过有效数字的位数可以判断近似值的精度：77%表示精确值在76.5%和77.5%之间，60%表示精确值在55%和65%之间，等等。在计算几个数的乘积时，最好的策略是使每个数都保持足够的精度，最后对计算结果取整。在某些场合，我用字母K表示1 000，也就是说，25K英镑表示25 000英镑。

你可以把这本书当做一个挑战。当你读完书中的问题之后，建议你不要直接看我的分析和结论，先独立地考虑一下。在某些章之后，我准备了练习题，目的是检验读者是否已经掌握了书中介绍的方法。在书末有答案。关于阅读顺序，建议先读第一章，在第十章以前先读第六章，在第十二章以前先读第八章，最后读第十三章。除此之外，你可以随意安排顺序。

本书不是系统化的数学教材。许多人的著作、演讲和谈话 viii 影响了我关于概率的理解，我不可能一一列出。本书中用到的技巧，除非特别声明，不是我首创的。在参考文献中，我仅列出了书中直接引用过的著作。我想向一些直接帮助过我的人致谢。牛津大学出版社的苏珊·哈里森（Susan Harrison）鼓励我写这本书，并通读了全文。吉夫·亚德兰（Geoff Adlam），约翰·巴雷特（John Barrett），特雷夫·波蒙特（Trevor Beaumont），麦克·博克特（Mike Birkett），唐比·布莱雷顿（Toby Brereton），约翰·克拉切尔（John Croucher），弗兰克·德克华兹（Frank Duckworth），约翰·维恩（John Gwynn），鲍勃·亨纳利（Bob Henery），雷顿·华封·威廉姆斯（Leighton Vaughan Williams），哈里·乔（Harry Joe），罗伯特·马修（Robert Matthews），哈格·莫

顿（Hugh Morton），路易丝·O·克纳（Louise O’s Connor），汉斯·利德维尔（Hans Riedwyl），乔纳森·西蒙（Jonathan Simon），伊安·沃克（Ian Walker），格雷格·伍得（Greg Wood）和比尔·辛巴（Bill Ziemba）为本书提供了宝贵的资料。本书关于大英帝国彩票的资料来自于网站 http://lottery.merseyworld.com/，感谢理查德·拉伊德（Richard Lloyd）对这个网站的维护和更新。我对数学知识最初的兴趣来自于我以前的老师，他们是伯特·劳治（Bert Lodge），詹福里·拉伊德（Jeffrey Lloyd），乔夫里·怀特海（Geoffrey Whitehouse），肯·西尼尔（Ken Senior），菲利斯·鲍曼（Phyllis Bowman），约翰·利维斯（John Lewis），大卫·肯道（David Kendall）和约翰·凯曼（John Kingman）。我的许多知识得自于我未曾谋面的人——威廉·费勒（William Feller），他的著作启发了一代概率学家的成长。亚当·黑格和丹尼尔·黑格（Adam and Daniel Haigh）教会我如何用家用电脑写作，凯·黑格（Kay Haigh）机智的评论也给我巨大的启发。

尽管我已尽最大的努力，我相信本书依然难免存在错误。希望这些错误不会构成对读者的误导。概率论是一个复杂的领域。如果您发现我在某些地方犯了原则性错误，请您告诉我。ix 在这个领域，比我强得多的数学家也犯过错误。

目　　录

第一章　什么是概率

无论概率是什么，你不能忽视它。当你向保险公司付保险费时，你的行为反映了保险公司付给你赔偿金的概率，以及赔偿金的数额。你是否应该为预防感冒注射疫苗？你必须考虑因注射疫苗而危害你的健康的可能性和因不注射疫苗而危害健康的可能性，并在二者之间做出权衡。如果你是一个刑事案件的陪审团成员，判决的依据是你相信被告的罪行是“没有疑问的”；而当你裁决民事案件时，相应的原则是“在各种可能性之间权衡”。当你买（或不买）大英帝国彩票时，你的行动可能仅仅出于兴趣或冲动，但是在一定程度上也依赖于你对获奖的可能性的估计。当你面临各种决定时，如果你对概率有深刻的理解，通常你能做出更好的决定。我的信条是有知识总比无知好，而获得关于概率的知识的主要途径在于研究那些依赖于机会的游戏。我的目标是展示一些主要的方法：如何评价你在游戏中获胜的可能性，以及在什么情况下你的机会可以增加。

精通概率本身不足以帮助你做出正确的决定。在很多情况下，你的知识只是帮助你在犯错误以后寻找失误的原因——事实如此。但是，知识总是有好处的，比较而言，如果你对概率

有更好的理解，你在生活和游戏中更容易做出好的决定。本书不是概率理论的教科书，它是一系列与概率相关的智力挑战。

语言的魅力在于同一个事物可以用不同的方法表达。在一副洗好的扑克牌中抽出最上面的一张[①]，你可以说："这张牌是黑桃的概率是四分之一。"下面三句话的意义与此完全相同：

1

1. 这张牌是黑桃的机会是四分之一。
2. 这张牌是黑桃的比率是一比三。
3. 这张牌不是黑桃的机会是这张牌是黑桃的机会的三倍。

另外一些词汇可以粗略地表达与"概率"相近的含义，比如"风险"、"几率"等等，相应地我们要涉及"预期"、"可能性"。然而，无论我们用什么词汇表达，我们希望表达的含义是什么？为什么我们要选择四分之一而非其它的数值？

"四分之一"这个数来源于一个理想化模型：一共有 52 张牌，其中 13 张是黑桃，在洗好牌后每一种可能分配出现的机会都是相等的——如果以上假设都正确，则这张牌是黑桃的概率是四分之一。以上表述并非绝对无懈可击，但已经可以足够精确地描述实际情况，而且不会引起误解。为了研究物理世界中的事件，不一定需要绝对完美的模型，即使在向月亮和火星发射火箭时，我们使用的模型也仅仅是近似的。

抽扑克牌的事件可以多次重复。为了检验我们的答案是否正确，我们可以把实验重复很多次，然后计算统计结果。如果重复的次数足够多，平均而言，结果是黑桃的情况将占全部实

① 译者注：西方人玩扑克牌的习惯是去掉王牌

验次数的四分之一。需要强调的是，我们没有说每抽四张牌就会恰好抽出一张黑桃，事实很可能是这样：我们连续好多次抽出黑桃，或者连续十几次都抽不出黑桃。这里有一个问题：我们刚刚用到的两个词——“平均”和“足够多”——是什么意思？100 次算不算足够多？1 万次呢？很不幸，这个问题没有简单的答案。

有时我们也讨论一次性事件的概率，比如明年股票市场的价格指数上涨 10% 以上的概率，或者巴西队夺取下一届世界杯足球赛冠军的概率。对于这种概率，无法用重复实验的方法检验，因为事件本身是不可重复的。关于概率的讨论也不一定限制为对未来事实的预测，比如，一个学者可以说“莎士比亚是《麦克白》的作者的概率是 80%”，这种关于概率的表达也是有意义的。

在以上各个例子中，概率都可以理解为对“相信的程度”的表达。我们对一个事件相信的程度越高，相应的概率值也就越大。如果我想表达我相信“股票市场价格指数将上涨一定数量”的程度，我可以和一个经纪人打赌，并且约定一定的赔率：比如说 1 赔 2。换句话说，如果我赌一块钱并且获胜，我将拿回 3 块钱（原有的 1 块钱加上赢得的 2 块钱）。此时，我的行为就反映了我对这件事的概率的估计。如果我觉得 1 赔 2 是对我有利的，就说明我认为其概率高于三分之一；如果我觉得 1 赔 2 是对我不利的，就说明我认为其概率低于三分之一。当然，你也可以就这件事打赌，并且你的结论很可能与我不同——这说明概率的值是因人而异的。在抽扑克牌的例子中，我们的观点是一致的，因为我们有一个基本一致的检验模型。在其它情况下——尤其是当我们掌握的信息不同时——我们的判断可能有巨大的差异。比如说，在一次赛马中，一匹马的训练

师和经纪人以及普通下注者的预期就不一样。

在涉及到掷骰子、抛硬币、赌扑克牌之类的例子时，存在一些已经被广泛接受的计算方法，大家很少产生分歧。大家可能出于不同的原因得出“某个轮盘赌是公平的”的结论，但是每个人的计算方法都是一样的。本书的多数内容将讨论这类情况。但是在某些情况下，专家的意见与普通人截然不同。去发现这种分歧是否对我们的决定或结论产生重要的影响，是很有价值的。

直观的考虑

维多利亚时代的学者弗兰西斯·高顿（Francis Galton）提出了一个容易算错的问题：如果抛三枚硬币，结果三枚硬币都是正面向上或都是反面向上的概率是多少？首先看一个马虎的计算：

在三枚硬币中，至少有两枚结果是一样的，或者都是正面向上，或者都是反面向上。而第三枚硬币出现正面和反面的概率是相等的，即它有二分之一的机会与另外两枚结果相同，同样有二分之一的机会与另外两枚结果不同。于是得出结论：三枚硬币都是正面向上或都是反面向上的概率是二分之一。

抓住以上论证中的错误需要逻辑技巧。一个办法是给三枚硬币编号，分别标为 1 号、2 号和 3 号，在排列所有可能结果时按编号顺序表示。通过这个办法可以避免高顿给出的错误结论。

你会发现，一共有 8 种可能结果，即

正正正，正正反，正反正，反正正，正反反，反正反，反

反正，反反反。

其中正好有两种结果符合要求。于是我们得出一个不同的答案：三枚硬币都是正面向上或都是反面向上的概率是八分之二，即四分之一。

高顿方法的错误在于一个意义模糊的词组——“第三枚硬币”。如果我们不对三枚硬币编号，我们怎么知道哪枚硬币是第三枚硬币呢？如果三枚硬币中恰好有两枚是正面，那么剩下的一枚就是第三枚硬币——可是这枚硬币必须是反面，高顿方法中的结论“第三枚硬币出现正面和反面的概率是相等的”是错误的。如果三枚硬币都是正面，那么无论我们把哪一枚当做第三枚硬币，第三枚硬币都应该是正面。无论如何，“第三枚硬币出现正面和反面的概率是相等的”这一结论是错误的。

错误的判断足以引起财产的损失。假设安迪知道正确的答案是四分之一，他愿意按 2 赔 1 的赔率赌三枚硬币的结果不同。他相信他将在赌博中获利：如果每次下注一块钱，在每四局他将赢三局，即赢三块钱，输两块钱。平均而言，在每四局中他将赢一块钱。假设波林相信高顿给出的错误结论，波林将很高兴地接受安迪的条件。他认为他有二分之一的机会输一块钱，但是有二分之一的机会赢两块钱，所以赌博对他有利。两个人都愿意赌，可是波林的分析是错误的——赌得越多，输得越惨。他应该反省自己的分析。

不变的法则

在表达概率的词汇中有一些描述程度的词，比如“不大可能”、“有可能”、“很可能”、“几乎确定”，等等。可是，当你

今天说某事“很可能”时，你的意思可能与昨天你用到这个词时不一样；同样，你所说的“很可能”与我所说的“很可能”也是不同的。为了达成一致，我们用数字表达概率。

表达概率的数字介于0和1之间。“不可能”用“概率为0”表示。在我看来，一个人穿越时空会见莫扎特的概率就是0；你可以用自己的理解说明“概率为0”的意义。另一方面，“概率为1”表示“必然”。我认为，猫王已死的概率就是1。有些人显然愿意打赌自己在10年之内不会死去，此时概率介于0和1之间。无论我们讨论的是事实还是猜想，概率的值必须介于0和1之间——这是不变的法则。

（尽管如此，我还是想介绍一段我的亲身经历。我刚刚做大学讲师的时候，曾给学生出过一道关于两个概率之间的关系的考题。我很漂亮地解决了那道题，答案看起来蛮有道理。然而，不幸的是我的答案中包括两个在0和1之外的概率值。幸亏我的一个同事发现了我的错误，使它没有流传出去。）

你可以用各种形式表示概率值，比率、分数、小数、百分比等等，取决于你的喜好。在各种结果出现的机会均等的情况下（比如掷骰子、抽扑克牌、玩轮盘赌等等），以分数表示概率值是很自然的。处理这类问题时，采用分数的表示形式通常会带来计算上的便利；而在计算的早期就把分数转换成小数通常是错误的。如果我们的目的是比较几个概率值，百分比、小数和公分母的分数更方便。

两个基本概念

附录2介绍了计算概率的方法，并给出了一些例子。有兴

趣的读者可以仔细研究附录2。不过如果你的目标就是读懂这本书，根本不必学习复杂的数学技巧，只要掌握两个基本概念就可以了。第一个基本概念是**排斥事件**：如果两个事件不可能同时发生，则这两个事件就是排斥事件。例如，在一副扑克牌中抽出一张，它可能是红心，也可能是方块，但不可能既是红心又是方块——这两件事是排斥的。考虑另一种情况，这张牌可能是草花，也可能是“K”，而且可能既是草花又是“K”——这两件事不是排斥的。如果两个事件a和b是排斥的，事件“a发生或者b发生”的概率就是a的概率与b的概率的和。还用刚才的例子，这张牌是红心的概率是四分之一，是方块的概率也是四分之一，于是这张牌是红心或者方块的概率是四分之一加四分之一，即二分之一。

第二个基本概念是**独立事件**：如果一个事件发生与否对另一个事件的概率不产生任何影响，则这两个事件是独立事件。例如，我抛一枚硬币，你也抛一枚硬币，这两枚硬币出现的结果可以认为是独立的。但是下一个例子与此不同：我从一副扑克牌中抽一张，然后你从剩下的牌中抽一张，抽出的结果不是独立的。假如我抽出了一张“A”，则在所有牌中“A”占的比例会下降（由原来的4/52变成3/51），所以你抽出一张“A”的概率已经受到影响。通常两件事是否独立是很明显的。关于独立事件的严格定义是：如果两个事件a和b同时发生的概率等于a发生的概率与b发生的概率的乘积，则这两个事件是独立事件。 5

把一个普通的骰子抛一次，得到“6”的概率是1/6。抛两次，两次都得到“6”的概率是（1/6）×（1/6）=1/36，这两个事件是独立的。换一个例子，假设大约每3年发生一次大型的海上石油泄露事件，而大约每4年伦敦队夺得一次英国

足球联赛的冠军，我们很难设想这两件事是相互影响的，所以有理由认为这是两个独立事件。于是，随机地选择一年，在这一年里发生了大型石油泄露并且伦敦队夺得了联赛冠军的概率是 (1/3) × (1/4) = 1/12。

在某些情况下，事件的独立性不容易发现。抛一下骰子，然后考虑两件事：

a：得到的结果是偶数。

b：得到的结果是 3 的倍数。

这两件事是独立的。就是说，假设我们已知骰子的结果是偶数，这对结果是不是 3 的倍数没有影响；反过来说，假设我们已知骰子的结果是 3 的倍数，这对结果是不是偶数也没有影响。在骰子的六个面中，每面朝上的概率都是 1/6；总共有三个面对应的数字是偶数，所以得到的结果是偶数的概率是 3/6 = 1/2（在这三个面中，具体某一个面向上是**排斥事件**，所以计算的方法是求三个概率的总和）。类似地，得到的结果是 3 的倍数的概率是 2/6 = 1/3，这两个概率的乘积是 1/6。另一方面，a 和 b 同时发生就意味着结果必须是 6，而这件事的概率也是 1/6。根据独立事件的定义，**如果两个事件 *a* 和 *b* 同时发生的概率等于 *a* 发生的概率与 *b* 发生的概率的乘积，则这两个事件是独立事件**，所以这两件事是独立的。

普通的骰子有六个面，此外也有一些特殊的骰子。取一枚金字塔型的骰子，它的形状是正四面体，各个面分别标上 1、2、3 和 4。把这个骰子抛一下，再考虑 a 和 b 的概率，此时这两件事就是不独立的了。如果已知骰子的结果是偶数，则结果必须是 2 或 4，所以它不可能是 3 的倍数；反过来，如果已知

骰子的结果是 3 的倍数，同样可以确定，结果是偶数的概率为 0。具体地说，事件 a 的概率是 1/2，因为在四个可能出现的数字中有两个是偶数；事件 b 的概率是 1/4，因为在四个可能出现的数字中只有一个是 3 的倍数；这两个概率的乘积是 1/2 × 1/4 = 1/8。然而，骰子的结果不可能既是偶数又是 3 的倍数，因为这个骰子只有 4 个面，没有数字 6。于是，a 和 b 同时发生的概率是 0，不等于 1/8。 6

如果把骰子换成钻石形骰子（正八面体，八个面分别标以 1 至 8），此时 a 和 b 又变成独立事件了！骰子的结果是偶数的概率是 1/2，是 3 的倍数的概率是 1/4。如果它既是偶数又是 3 的倍数，则它必须是 6 的倍数，在 8 种可能结果中只有一种符合条件，所以 a 和 b 同时发生的概率是 1/8。由于（1/2）×（1/4）= 1/8，依据独立事件的定义，a 和 b 是独立的。

在此例中，如果已知结果是 3 的倍数，则结果只能是 3 或 6。在两个可能结果（3 和 6）中有一个是偶数（6），所以在此前提下结果是偶数的概率是 1/2。因此，无论事件 b 是否已经发生，事件 a 的概率都是 1/2。同样，如果已知结果是偶数，则结果可能是 2、4、6 或 8。在四个可能结果中只有一个是 3 的倍数（6）。所以，无论事件 a 是否已经发生，事件 b 的概率都是 1/4。

在以上各例中，事件 a 和 b 是否独立取决于骰子有几个面。在具体的问题中，需要全面地分析。如果你希望借助直觉判断两个事件是否独立，只需考虑一个问题：假设你已经知道其中的一件事已发生，另一件事的概率是否因此改变？如果不改变，则两件事是独立的。

多次重复的事件

虽然我强调概率所表达的是你个人对某件事相信的程度，它对于预测多次重复的事件的结果还是有用的，比如抛硬币的例子。抛一次硬币，出现正面的概率是 1/2，如果抛 1 000 次或 100 万次，我们对结果的预测应当如何？我们是否应当期望出现正面的比例正好是 1/2？你可能漫不经心地回答“是”，而这恰恰是需要深思的。

抛 100 万次硬币，而期望正面和反面都恰好出现 50 万次，这是没道理的。如果你期望出现正面的比例落在包含 1/2 的一个固定的区间内，那就是另外一回事了。假设你期望出现正面的比例在 49%至 51%之间，而总共抛 100 次硬币，则你的要求是正面出现的次数是 49、50 或 51。你可以实际检验一下，做 100 次实验，符合要求的结果所占的比例大约是 24%。[①]

下面把实验的次数增加到 1 000 次，则正面出现 490 次至 510 次都可以是符合要求的，此时，符合要求的结果所占的比例大约是 50%。当实验的次数增加到 1 万次时，符合要求的区间是 4 900 至 5 100，而符合要求的结果所占的比例高于 95%。而当实验的次数增加到 100 万次时，出现正面的比例落在 49%至 51%之间的概率已非常接近 100%。实际上，出现正

① 译者注：做 100 次检验，每次检验包括 100 次抛硬币的实验，根据概率可以期望 100 次检验中大约有 24 次符合要求——正面出现的比例在 49%至 51%之间，这就是概率 24%的含义。下面讨论的几种概率与此类似，作者想要证明的是当实验次数增加时根据概率做出的预期的可靠性上升。

面的比例落在 49% 至 51% 之外的概率比日丹诺夫重新当上俄罗斯的领导人的概率还低。

即使我们把目标区间限定得更严格，结论依然是相似的。假设我们要求出现正面的比例在 49.9% 至 50.1% 的区间内，当实验次数达到 1 000 时符合要求的概率是比较低的（这是要求出现正面的次数恰好是 499、500 或 501），而当实验次数达到 1 000 万次时符合要求的概率非常接近于 100%。最终，正面出现的比例将接近 1/2。需要注意的是，你不能指望正面出现的次数与反面出现的次数完全相等，或者二者的差小于一个固定的数（比如说 20）。事实上，符合这种要求的概率是极低的。如果实验的次数足够多，就概率而言，正面出现的次数与反面出现的次数会有很大的差距，这个差距可以超过任何你指定的数。在这个过程中，实际上保持稳定的是正面出现的次数占全部次数的**比例**。

在讨论的期望值不是 1/2 时以上结论也是有效的。在一副洗好的扑克牌中抽一张，抽出黑桃的概率是四分之一，即 25%。如果重复实验很多次，抽出黑桃的次数在 24% 至 26% 之间的概率是多少？与抛硬币的例子一样，最终的结果依赖于实验的次数。如果只抽 10 次，则符合要求的概率是 0（因为 24% 至 26% 意味着 2.4 至 2.6 张黑桃被抽出，这是不可能的）。如果抽 100 次，则要求抽出黑桃的次数是 24、25 或 26，此时概率大约是 1/4。而如果抽 100 万次，符合要求的范围是相当大的：从 24 万到 25 万都符合要求，此时，出现黑桃的次数落在这个范围以外的可能性极低——与一匹谢特兰群岛的小马驹在大英帝国马赛上夺得冠军的可能性相仿。

下面对以上情况做一个总结。对一个可重复性的事件而言，事件可以在相同的背景下重复发生，而且每次事件是独立

的。我们所要求的情况出现的概率是一个固定值 x，多次重复此事件，并跟踪研究我们所要求的情况实际出现的频率。最初这个频率在 x 周围震荡，而且与 x 的偏差可能很大，但是，我们所关心的是事件多次重复以后的结果。我们设定一个包含 x 的很小的固定区域（比如在抛硬币的例子中设定一个包含 1/2 的小区域——49.9% ~ 50.1% 或 49% ~ 50%），观察实际出现的频率是否落在此区域内。当实验次数较小时频率经常落在区域之外，但是当实验次数足够多时，频率将落在区域之内，并且随着实验次数的增加保持在此区域内。如果存在所谓的“平均律”，它就是这样起作用的。只有在事件多次发生的情况下它才发生作用，对于短期行为它不能做出任何预测。当我们谈到“平均律”的时候，“平均律”的意义总是依赖于具体语境的，不精确的。在抛硬币的例子中，也许你会得出结论：“由于出现正面的概率与出现反面的概率相等，所以下一次抛硬币的结果将使出现正面的频率与出现反面的频率趋于相等。也就是说，如果到目前为止正面出现的次数多于反面出现的次数，则下一次出现反面的概率高于出现正面的概率。”这是错误的！平均律不会告诉我们任何关于下一次（或下 100 次）事件的结果。根据平均律我们确实可以计算出现某种结果的频率，但是仅仅就“最终”的结果而言。

如果我们可以按这种方式重复实验，多次重复的结果可以为我们提供估计概率值的根据。我们不断计算某种结果出现的频率，并且保证实验的次数足够多，则这个频率就可以作为表达概率的依据。我们得到的估计值有时候会比实际的概率高，有时候会比实际的概率低，但平均而言结果是正确的。重复的次数越多，结果的可靠性越好。

有些事件是一次性的。一个天气预报员可以预言“明天下

雨的概率是50%"。她可以通过分析气压图、卫星数据以及历年同期的天气状况等资料得出结论，她有一些预测天气的模型。然而，在抛硬币的例子中使用的检验方法对她来说是无效的。"明天"是一个特定的日子，它要么下雨，要么不下雨。她永远无法确知她的预言是否准确。为了检验她的预言，她可以积累几年以来她的预言的纪录。其中可能有100次她预言下雨的概率是50%；另外有80次她预言下雨的概率是25%，等等。按这种方式把她的预言分组，比较预言与实际的天气状况，此时情况还原为可多次重复的事件的情况。用这种方法可以检验一次性事件的概率——当然，我们需要假设她的预测能力是保持稳定的。与可重复性事件完全相同，对于每一个单独的预言我们无法判断，但是我们可以检验她的预言的整体的正确性。

平均值与变异性

在许多学科中，最有价值的观念往往是最简单的观念，在概率论和统计学中也是如此。当你玩一个游戏时，你可能知道你将赢得的数额的各个可能值，以及每个值出现的可能性有多大，但是真正决定这个游戏对你是否有利的是你的收入的平均值。在统计学的领域内，为了做出合理的决定，平均值往往比许多单独的数值更有用。本书假设读者已经在生活实践中形成了关于"平均值"的直观理解：你知道你的普通旅行所占的平均时间，你在超级市场的账单的平均数额，你在各种游戏中获胜的平均频率，等等。在多数场合，这种直观对于我们已经足够了，但有时我们需要精确的计算。附录3中介绍了严格的计

9

算方法。

考虑这个例子：联盟杯足球赛即将举行，16 支球队分 8 组比赛。分组情况尚未公布，你的任务是猜测分组的结果，并且使你的预测与实际情况尽量相符。我们不关心球队的主客场，只关心你的配对方案与实际是否相符。你每猜对一组配对，我将付给你 1.5 英镑，但是为了与我打赌你必须先付给我一笔赌注。你愿意付多少钱？在继续阅读之前，请计算这个问题，并记下你的答案。

你最多可能赢得 12 英镑，如果 8 组配对你都猜对的话；当然，你也可能一无所获。如果你想计算出你可能赢得的各种数额，以及每种数额对应的可能性，你会发现这个过程相当繁琐。然而，我们不需要了解计算的细节。如果不考虑参与本身带来的乐趣，左右你的决策的是你将赢得的平均数额。如果你的赌注比这个平均数额低，则这个赌博对你有利；否则对你不利。

如果你精于概率计算，你将发现你将赢得的平均数额是 0.8 英镑，所以，当赌注低于 0.8 英镑时游戏是对你有利的。附录 3 给出了计算方法。需要强调的是，为了做出决定，我们必须忽略对各种可能结果的细节的考虑，而坚持一个简单的判定标准。任何不高于 0.8 英镑的赌注你都乐于接受，而任何不低于 0.8 英镑的赌注我都乐于接受。0.8 英镑的数额将导致一场公平的游戏，任何一方都不能期望从中获利。

下面改变游戏规则。我不再为每一组正确的配对付款，而是要求你最少猜对一定数量的配对。如果你没有达到要求，你将一无所获。以下是三种游戏规则（最低要求分别是 2、4 和 8）：

（1）如果你猜对至少 2 组，你赢得 7.94 英镑；

（2）如果你猜对至少 4 组，你赢得 344.66 英镑；

（3）如果你猜对全部8组，你赢得16 216 209英镑。

（补充一点，假设我有能力支付这笔巨款。）

在这三个游戏中，你将赢得的平均数额都是0.8英镑，所以要求你支付0.8英镑的赌注是公平的。与最初的例子相比，在这三个游戏中你将赢得的平均数额是一样的，但是这四个游戏对你的吸引力是不同的。一个根本的差别在于变异性。

在最初的例子中，每次你有40%的机会获胜；在另外3个游戏中，你获胜的机会分别是10%、2%和200万分之一。一个谨慎的人可能更乐于参加最初的游戏，因为在这个游戏中多次下注而一无所获的可能性是很小的。如果你不那么谨慎，最后一个游戏可能对你更有吸引力。你可能不在乎输掉许多个0.8英镑，而寄希望于获得巨奖，虽然希望是如此渺茫。

实际上最后一个例子与保险非常相似。当你驾驶一辆汽车时，你可能撞坏东西或撞到人（保险业所谓的"第三者"），你可能因此得付巨额赔偿。绝大多数人没有能力支付这样一笔巨款，所以我们向保险公司投保。保险费通常涵盖许多其他的风险，当我们投保的事件确实发生时，保险公司替我们支付100万英镑的赔偿金。保险公司的估计可能是每年这种事件发生的概率是10万分之一，并且向我们收20英镑的保险金。

我们把这个例子当做一个赌博游戏分析。你付出20英镑，在绝大多数情况下，你什么也赢不到。有10万分之一的机会你可以"赢"100万英镑（保险公司替你支付100万英镑的赔偿金）。你在这个游戏中将赢得的平均金额是10英镑。单就数额而言，这个游戏是对你不利的。但是，如果你是一个谨慎的人，你会认为这个游戏是对你有好处的！你真正关心的不是平

均支出，而是你确实有可能陷入拼命寻找100万英镑的窘境，事先付出的20英镑保险金是安全保证。这种游戏可以安慰你的心理：你的安全受到了保护。

在这个游戏中保险公司也感到满意，如果它能找到足够多的客户。如果客户的基数达到1 000万，每个客户付20英镑，则每年保险公司收入两亿英镑。由于每个客户索赔的概率是10万分之一，所以平均而言有100个客户索赔，赔偿金总额为1亿英镑。因此，公司可以期望可观的利润。但是，如果有200个以上客户索赔，公司将遭受损失。这种情况发生的概率有多大？保险公司知道，在这种客户规模下危险几乎不存在。平均而言会有100个客户索赔，如果实际索赔的人数达到110或115，这并不奇怪；可是，如果某一年索赔人数达到150或11 更多，其可能性低于千年一遇。

对于客户数只有1 000人的小保险公司而言，情况则截然不同。此时公司一年的收入为200万英镑，平均而言将有一个100万的索赔发生。这个比例与大保险公司相同，但是每12年将出现一次索赔人数达到3或更多的情况。一旦这种情况发生，支出将超过保险金。12年1次的频率是令人不安的，对于公司和客户双方这都是潜在的灾难。对于这个小保险公司而言，可能支出的变异性太大了。保险公司的经营策略是通过增加客户数降低破产的风险。

这是一个重要的普遍原则。如果游戏只进行一次，了解平均值是必须的，但仅仅考虑平均值是不够的。当游戏仅进行一次或几次时，支出的变异性可能比平均值更重要；当游戏的次数非常大时，则平均值成为主导因素。保险公司的情况就是和非常多的客户同时玩一个游戏。为了使平均值成为主导因素，12 支出的变异性越大，需要的游戏次数就越多。

第二章　彩　　票

总有人请教数学家买彩票的制胜战略。他们的目的是希望数学家提供一种预测中奖号码的数学方法。数学家莱恩·夏克莱顿（Len Shackleton，我童年时崇拜的偶像）的自传中有一章专门讨论“一个普通的足球教练对足球了解多少”，这一章仅仅是一页白纸。如果你要求我写一章“如何通过彩票赢更多的钱”，恐怕我也只能交白卷。我所能提供的惟一建议就是“买更多的彩票”。虽然如此，数学方法仍然可以指导你如何买彩票。下面我就要介绍这些方法。

始于1994年9月的大英帝国彩票。在商业上它是个巨大的成功。英国超过90%的成年人买过这种彩票，通常每周销售额超过8 000万英镑。它的标识已广为人知，那句响亮的广告词“下一个赢家是你！”始终展示着头奖的诱惑。大英帝国彩票已经造就了数以百计的百万富翁，头奖奖金超过2 200万英镑。

世界各地流行着类似的游戏。为了表述完整，我仔细介绍一下英国的规则。要加入一轮游戏，你需要花1英镑买1张彩票，并且在1至49之间选六个号码。这叫做“49分之6号码

游戏"。在每一轮，由一个特制的机器从 49 个小球中选出 6 个，每个小球代表一个号码。如果你挑选的小球中至少有 3 个与机器跳出的相符，你就中奖了。相符的小球越多，你得的奖金就越多。如果六个小球你都选中了，你就赢得头奖。

除了六个小球都选中以外，你还有一个额外机会。机器还选出第七个号码（特别号码），如果你选中的六个号码中有五个相符，而第六个号码与特别号码相符，则你赢得次奖。否则，特别号码无效。

彩票的组织者保证每个号码都是随机产生的，作弊是不可能的。每个小球被尽量制造成完全一样的，尺寸、重量、湿度、硬度和表面光滑程度都没有差别。共有八组编好号的这样的小球和 4 台不同的机器。每台机器实际上是一个透明的塑料桶，其功能是彻底打乱小球的顺序。在开奖的当天，八组小球中的某一组和四台机器中的某一台被挑选出来。抽奖的过程在电视上直播，并且有一位普通公民被邀请到现场作见证。每次抽奖结束后，机器被倒空，所有 49 枚小球被装入一个盒子里，贴上封条——这个过程要在一个独立的审计员的监视下完成。这些小球在下一次使用以前要经过度量衡标准局的检查。在两次抽奖之间，所有的器材被装在一个密封的容器内。

概率论和统计学可以帮助我们解决三个问题：

○ 你获奖的机会有多大？

○ 如果你参与，如何提高获胜的机会？

○ 如何辨别机器产生的号码是不是随机的？

你获奖的机会

大英帝国彩票的吸引力在于数额巨大的头奖。许多人买彩票的原因是他们仅仅关注一个事实：天文数字的巨额奖金确实颁发过许多次。其实，一个合理的分析不仅应当关注奖金的数额，同时应当考虑获奖的机会有多大。

当你看到小球从机器中滚出来时，很容易相信每一个小球第一个滚出来的概率是一样的。也许这个小球是16号。接下来，剩下的48个小球中的每一个成为第二个的概率也是一样的，以此类推。当六个小球全部产生以后，我们发现任何一组六个号码的组合出现的概率都是相同的。也就是说，{1，2，3，4，5，6}出现的概率与{3，4，20，38，48，39}出现的概率相等，虽然后者显得更加随机。如果你对此有疑问（许多人确实如此），你就是相信有一些组合出现的概率天然地高于其它组合。也许你应该把你认为出现的概率低于平均值的组合列成一张表，而后校正你的选择。

借助于对称性可以证明我们的结论。在我看来，没有任何理由相信某一个组合出现的概率可以比其它组合高或低。然而，我在长期的实践中得到一个知识：如果一个人相信{1，2，3，4，5，6}出现的概率比其它组合低，通过论证让他改变观点几乎是不可能的。如果你在酒馆里遇到这样一个顽固的 14 家伙，我建议你用实战改变他的立场。在台球桌上拿5个台球（1号至5号），放在一个书包里，让他从书包里随意地摸出3个球，则一共有10种可能组合（附录1中有证明）。由于在大英帝国彩票中他相信{1，2，3，4，5，6}出现的概率低于平

均值，同理现在他也应当相信 {1，2，3} 出现的概率低于平均值。你和他赌 {1，2，3} 将出现，他应当给你高于 1 赔 10 的赔率。和他打这个赌，如果你不能改变他的观点，至少你可以改变自己的经济状况。

在这个问题上统一观点以后，我们还需要计算一共有多少种可能的组合。附录 1 中给出了计算方法，答案是 ${}^{49}C_6$，它的值是 13 983 816。

只买一张彩票，六个号码全对的机会是 13 983 816 分之 1，大概相当于 1 400 万分之一。我们如何理解这个数字呢？如果换成二分之一或者十分之一，我们倒容易理解，在日常生活中经常用到这样的数字，比如在预测一次抛硬币的结果或估计你的汽车不能通过安全检查的可能性时。可是当概率值变成 100 万分之一或者 10 亿分之一时，我们得到的信息只是“其可能性小得可以忽略不计”。这两个数字对于我们没什么差别，虽然前者是后者的 1 000 倍。

假设你每周花 1 000 英镑买彩票，平均而言，每 270 年将有一次六个号码全对。换一个说法，假设你的祖上从 5 000 年以前开始买这种彩票，坚持每周买 50 张，假如运气一般，那么到今天你的家族将赢得一次大奖。为了这一次幸运的降临，你的家族必须在辽远的时空中经历漫长的等待。当然，在等待的过程中，你的家族会赢得许多次小奖，但是对每一张彩票而言，六个号码全对的概率是非常非常低的。

为了计算赢得数额较低的奖金的概率，我们需要考虑有多少张彩票恰好选中了三个号码，以及有多少张彩票恰好选中了四个号码，有多少张彩票恰好选中了五个号码。我们把机器最终跳出的六个号码称为“好号码”，而把其余的 43 个号码称为

15 “坏号码”。（以后我们再处理“额外机会号码”。）我们想知道

的是：在我们选出的六个号码中，有几个是“好号码”？有几个是“坏号码”？这是一个标准的排列组合问题，附录1中介绍了完整的解法。

首先考虑恰好选中4个号码的情况。由于恰好选中了6个“好号码”中的4个，所以同时也选中了43个“坏号码”中的2个。因此，恰好选中4个号码的彩票的总数是（$^{6}C_{4}$）×（$^{43}C_{2}$），它的值是13 545。用同样方法可以计算选中的号码数恰好为0、1、2、3或5的彩票总数，结果见表2.1。

表2.1：选中的号码数与对应的彩票数

选中的号码数	对应的彩票数
6	1
5	258
4	13 545
3	246 820
2	1 851 150
1	5 775 588
0	6 096 454

为了检验表2-1是否正确，你可以把表右侧的七个数字加起来，看看总和与彩票的总数是否相等。在计算中，我们没有假定具体哪六个号码是“好号码”，所以无论机器最终选出的号码是什么，以上答案都是有效的。

假定每种结果出现的可能性都是相同的，则表2.1已给出了计算各个概率值的全部信息，只要把对应的彩票数除以

1 400万就可以了。例如，恰好选中 4 个号码的概率就是 13 545/13 983 816，这个数略小于 1/1 000。如果你每周花 5 英镑买彩票，根据概率每四年你可以得到一次这样的机会。

我想你已经注意到了，在表 2.1 中最大的数出现在最后一行。这意味着可能性最大的结果是你一个号码也没选中，可能性第二大的是你仅选中了一个号码。所以，如果你只选中了一个号码，这并不意味着坏运气。要知道，有 85% 的彩票与你运气相当或者更糟。

16

我们可以很容易地算出平均每张彩票将选对几个号码。你选出的六个号码中的任何一个都有 6/49 的概率选对，所以平均而言，你将选对 6×（6/49）=36/49 个号码。（有些人可能对这个结论产生疑问，因为 36/49 不是整数。确实如此，但作为平均数它是有意义的，就好像我可以说你平均每周买 1.5 张 CD 唱片。）

表 2.2：各种情况对应的概率值

获奖情况	对应的彩票数	概率值
头奖	1	1/1 400 万
额外机会奖	6	1/230 万
选对五个号码	252	1/5 550
选对四个号码	13 545	1/1 030
选对三个号码	246 820	1/57
其它情况	260 624	1/54

现在我们计算获得“额外机会奖”的概率，我们将要考虑的情况是你选对了六个“好号码”中的五个而你选中的第六个

号码与“额外机会号码”相符。这个计算很简单。“额外机会号码”可以取代六个“好号码”中的任何一个，所以一共有六种赢得“额外机会奖”的可能组合。在表 2.1 中，有 258 张彩票恰好选对了五个号码，其中的六张获“额外机会奖”，其余的 252 张则否。表 2.2 给出了各种情况对应的概率值。

表 2.2 介绍了各种可能性。大英帝国彩票的设计者充分研究了各种类似的抽彩活动，并且调查了顾客购买彩票时的心理。彩票的设计依赖于两条基本原则：第一，每张彩票都有赢得奖金的合理机会，然而机会相当小；第二，赢得巨额奖金的机会是看得见、摸得着的。如果一个人每周花 5 英镑买彩票，平均每年他将中 5 次奖，这个频率已足以使他保持参与热情。而如果你每周花 1 英镑买彩票，赢得头奖对你来说始终是一个遥远的梦想（仅仅是梦想）。正是由于赢得头奖的机会是如此的小，所以头奖的金额可以非常巨大。抽彩活动定期地创造一夜暴富的神话，而每一个新生的百万富翁都被媒体大肆宣传，这正好进一步刺激了下一轮彩票的销售。然而，我们可以从中赢得多少钱呢？

在全部赌资中只有 45% 作为奖金发还。这意味着无论奖金的比例如何分配，也无论总共卖出了多少张彩票，平均而言你每付出 1 块钱将赢回 0.45 块钱。从这个平均值出发，这个游戏是不划算的。然而它的特点是，有 98% 的可能性你一无所获，在其他的可能性中，你将赢一大笔钱。巨额奖金的吸引力足以战胜对低可能性的不满，参与者往往被这个微弱的可能性所吸引。售出的彩票越多，奖金额越高。为了便于讨论，我们假设卖出了 1 亿张彩票，则奖金额为 4 500 万英镑。 17

在获奖号码产生以后，首先统计出所有恰好猜对了三个号码的彩票，每张彩票获 10 英镑奖金。依据概率，这笔奖金的

总额将占全部奖金的 39%。当然，实际比例有可能比 39% 高很多或低很多，在对奖活动结束以前我们无法确知精确值。为了计算简便，假设结果恰好等于平均值，即 1 800 万英镑。奖金中还剩下 2 700 万英镑，这笔钱称为“锅底”。

“锅底”被按照固定的比例分成 4 份，分别用来支付 4 个等级的奖金，其中最大的一份分给头奖。在我们的例子中，头奖的奖金大约是 1 400 万。具体某个获奖者赢得多少奖金依赖于有多少人与他同时获得该级别的奖金，因为 4 份奖金中的每一份都将被共同获奖者瓜分。如果只有一个人获头奖，则此人独得 1 400 万；如果有 10 人同时获头奖，则每人得 140 万；如果有 100 人同时获头奖，则每人的奖金只有 14 万。

你认为这种情况不可能发生吗？在 1995 年 1 月开出的第九轮彩票中，头奖的奖金达到 1 600 万英镑。就在一个月以前，布莱克波恩（Blackburn）的一个工人在第四轮彩票中独得头奖，拿走了 1 700 万英镑。可是在第九轮中，有 133 个人同时赢得头奖，每人的奖金大约是 122 510 英镑。这依然是一个可观的数字，可是与 1 700 万英镑相比是多么令人失望！如果你希望赢得一笔巨额奖金，你应当避免选择那些会与许多人重复的号码组合。下文将告诉你具体的策略。

如果一张彩票恰好选对了三个号码，奖金额是固定的：只有 10 英镑。对于其他等级的奖金而言，奖金额是不确定的。然而，我们可以计算出各个等级的奖金的平均值。在一轮正常的彩票中，无论实际上售出了多少张彩票，这个平均值是保持不变的。当然我们也会发现，奖金的实际值完全可能偏离平均值很远。

有时头奖的奖金额相当巨大。通常这是因为在上一轮抽奖中没有人获头奖，所以头奖的奖金累积到下一轮；还有一个可

能是彩票的发行者从储备金中拿出了一部分追加到头奖奖金中，这种情况称为“超级彩票”。无论在哪种情况下，其它等级的奖金不受影响。此外，制定抽彩规则的人注意到一种纯粹理论上的可能性：选中三个号码的彩票可能非常多，以至于为每张末奖彩票付 10 英镑就将耗尽全部奖金。为此，抽奖规则特别规定，在这种情况下，所有获奖彩票（包括头奖）平分全部奖金——也许是 8 英镑。当然，理智健全的人不会对这种可能性当真。 18

表 2.3 列出了各个等级的奖金的平均值，以及截至 1998 年 9 月为止已举办的 282 轮抽彩活动中产生的极限值（最大值和最小值）。数额最大的头奖（2 260 万英镑）是由累积奖金造就的。在没有累积奖金的情况下产生的最高奖金是 1 170 万英镑。有趣的是，数额最小的那笔头奖奖金同样产生于有累积奖金的一轮，即前文介绍过的第九轮。如果没有累积奖金，头奖奖金甚至会降到 48 500 英镑。头奖奖金的总额最高曾达到 4 200万英镑，那笔奖金最终被三张彩票平分。但是，假如那一轮的头奖号码组合是 {1, 2, 3, 4, 5, 6}，头奖奖金将跌到惨不忍睹的 1 400 英镑，因为据彩票的发行者介绍，那一轮有 3 万张彩票买了这个号码组合。

表 2.3：过去的 282 轮抽彩中的奖金额度

获奖等级	平均值	最小值	最大值
头奖	200 万	122 500	2 260 万
额外机会奖	102 000	7 900	796 000
5 个正确号码奖	1 520	181	7 800
4 个正确号码奖	62	16	164

实际的奖金额偏离平均值如此之远令人震惊。除奖金的平均值以外，这种变异性是最引人注目的现象。获奖当然比不获强，但是如果你只选中了5个号码，你的奖金还是不够买一辆新福特车。通过历史我们发现，如果你希望赢得一笔1万英镑以上的奖金，你必须中头奖或额外机会奖。以一张彩票赢得一笔这样的奖金的概率是200万分之一。

奖金数额变化如此之大是因为有一些号码组合比其它号码组合出现的频率高得多。在每一轮彩票中，发行者都确切地知道有多少张彩票买了某个号码组合，但是他们严守机密。据我所知，他们只公布过两条相关信息：

(1) {1，2，3，4，5，6} 出现的频率极高，超过每周1万张；

(2) {7，14，21，28，35，42} 创下了单轮售出数最高的纪录。

乔纳杉·西蒙报道了1996年10月举行的第101轮彩票的情况。在那一轮，平均每一个号码组合被买了5次。在开奖以前，头奖奖金的预期值达1 000万英镑。有2 000个号码组合吸引了大批的买家。如果这2 000个号码中的某一个中头奖，获奖的彩票数将超过200张，每个中奖者的奖金将低于5万英镑。这2 000个号码组合在全部可能选择中仅占7 000万分之一的比例，根据概率，如果每周开两次奖，在70年里这2 000个号码组合中将有一个获头奖。另外，有20万个号码组合的购买者超过了20人。

这一轮彩票中存在相反的极端。有三分之一的号码组合的购买者不超过两人。事实上，将近70万个号码组合（约占总数的5%）根本没人买，其中包括大家梦寐以求的头奖号码组

合。大约12%的号码组合仅卖出一张，另外16%的号码组合恰好卖出了两张。很明显，如果这样的彩票中奖，奖金将是500万英镑或更多。

许多人从抽彩活动一开始就买了一个号码组合，并且在以后的每一轮彩票中坚持买同一个号码组合，至今不敢放弃。参与者也可以选择买一张季票，即一次性地在许多轮彩票中买同一个号码组合。季票可以保证你的选择在每一轮彩票中保持稳定性，相反，你也可以每次都随机地选择号码组合，一种叫做“幸运星”的小玩艺（Lucky Dip），可以帮助你产生随机号码。

买彩票的人希望知道哪些号码组合是“流行”的，目的是避免和许多人同时购买一个号码组合。当然，如果有人发现了一个很少有人购买的号码组合，并且公布于世，大家一定会蜂拥而上买这个号码组合，从而把它变成一个“流行”的号码。我们是否有可能计算出哪些号码是“不流行”的呢？

获胜策略？

我将给出几个战略性的建议。它们主要告诉你什么错误应当避免，给出某些警告。大家购买彩票的行为应当做一些调整。如果你买彩票，好得很，希望你从中获得快乐，但是你最好别指望因此发家。

有些国家公开发布彩票的统计数字。在加拿大，每周公布各个号码被购买的次数。每次公布的数字会有一些变化，但是最流行的号码与最不流行的号码基本上保持不变。把所有号码按流行程度排列，大致是{7，11，3，9，5，27，31，8，17……30，46，38，45，20，41，48，39，40}。每周这个排行榜

会有微小的调整。表中未列出的号码被购买的频率居中。

根据这个排行榜你可能会认为{40，39，48，41，20，45}的号码组合是非常不流行的——大错特错。由于加拿大的彩票购买者已经知道这六个号码是最不流行的，所以好多人专门买了这个号码组合！（这是表现事件的独立性的绝妙的例子。六个最不流行的号码构成了一个最流行的号码组合！）常见的号码未必构成常见的号码组合，罕见的号码也未必构成罕见的号码组合。

瑞士彩票的规则与大英帝国彩票略有不同，它的方法是从45个号码中选取6个。根据附录1介绍的计算方法，在瑞士彩票中全部可选择的号码组合总数为$^{45}C_6 = 8\ 145\ 060$。汉斯·里德威公布了某一轮瑞士彩票中的有趣的统计数字。在那一轮中，一共卖出了16 862 596张彩票，平均每个号码组合卖出了2次。汉斯·里德威指出，至少有5 000个号码组合卖出了50次以上！最初看到这个统计数字时，我非常震惊。我渴望接触大英帝国彩票的统计数字，看看英国的彩票购买者是否干过类似的事。前文介绍过英国彩票第101轮的一些统计结果，我猜想当时一定发生了与瑞士相似的情况。汉斯·里德威的统计数字值得进一步研究。

瑞士彩票的外观如图2－1。45个号码排成7行6列，外加一个半行。你是否相信，有两个号码组合的购买者超过了2.4万人！这两个号码组合是由从左上到右下的对角线以及从右上到左下的对角线组成的，即{1,8,15,22,29,36}和{6,11,16,21,26,31}。另外，每一个整行构成的号码组合卖出了2 500次以上，许多同一个方格内的号码构成的号码组合（比如{8,9,14,15,20,21}）以及某些对角线号码构成的号码组合（比如{7,14,21,28,35,42}）也卖出了2 500次以上。整洁的几何模式吸引了

许多购买者。(如图 2－1)

图 2－1 瑞士彩票的外观

1	2	3	4	5	6
7	8	9	10	11	12
13	14	15	16	17	18
19	20	21	22	23	24
25	26	27	28	29	30
31	32	33	34	35	36
37	38	39	40	41	42
	43	44	45		

21

为什么有一个明显随机的号码组合{8，14，17，25，31，36}也卖出了 1.2 万次以上呢？因为这是上周的获奖号码！{9，15，18，26，32，37}和{7，13，16，24，30，35}也卖出了 4 000 次以上。你注意到了吗？那个获奖号码的各个数字加 1 或减 1 就得到这两个号码组合。每周都有数以百计的人追逐上一周的获奖号码。几年以来的获奖号码以及最近在奥地利、法国和德国获奖的号码也非常受欢迎。简略地说，如果有一种系统的方法产生六个号码，好多瑞士人会选用这种方法。

这是一个非常普遍的现象。在爱尔兰彩票中，有两个人平分了头奖。这两个人选择号码的依据是同一个牧师出生、任职和去世的日期。当罗马教皇约翰 13 世去世时，许多意大利人根据约翰死亡的时间、日期以及年龄选择自己的幸运号码。如

果你故意按照某种逻辑程序选择你的号码，你就必须冒一个风险：说不定在 3 000 万个购买者中有好多人和你打同样的主意！为了避免这种情况，你应当完全随机地选择你的号码。如果你自己都不知道将要选择什么号码，别人当然也不会知道。当然，如果你“随机选择”的结果恰好是{1，2，3，4，5，6}或者其它明显比较流行的号码组合，你最好放弃这个号码组合，重新选一次。

胡曼斯发现，如果不借助于骰子、硬币、扑克牌之类的工具，选择真正随机的号码是一件很困难的事。如果我们要求一些人完全凭想象力给出几个随机的数字和数字组合，在他们的答案中，将有一些数字和数字组合出现的频率很高，同时有一些和数字组合出现的频率很低。在本节的开头我们给出了加拿大彩票的号码频率排行榜，在那个排行榜中我们可以发现彩票购买者的心理。在排行榜中像 20、30、40 这样比较“整”的数出现频率很低，这说明参与者在潜意识中的一个期望：选择结果最好看起来像是随机的。你可以做一个简单的试验：让你的一些朋友随机地在 1 到 10 之间选择两个数字。除非他们看过这本书，否则他们极少选择 5，而且在绝大多数情况下，他们选择的两个数的差即不太大，也不太小，通常是 3，4 或 5。

新西兰彩票从 1987 年开始发行，抽奖规则是从 40 个号码中选 6 个，每周开奖。最近，新西兰彩票每轮销售量大约是 1 400万张。全部可能的号码组合大约有 400 万种，平均每个号码组合吸引了 3.7 个购买者。与英国人相比，新西兰人更热衷于使用“幸运星”之类的辅助工具生成完全随机的号码。在英国，大家渐渐接受了幸运星之类的工具，大约有 15～20%的彩票是用这种方法选择的，在存在累积奖金的情况下这个比例会更大一点。而在新西兰，62%的彩票是借助于生成随机号

码的工具选择的，只有38%的彩票是根据购买者的个人偏好选择的：生日，幸运号码，以前的获奖号码等等常常是他们追逐的目标。大量使用辅助工具的结果是每周获奖的人数大致稳定，所以奖金的数额更容易预测。如果购买者更多地随机选择号码，结果就更加稳定——粗看起来有点奇怪吧？可是事实就是这样：更多的随机选择带来了更好的可预测性！ 22

以上情况可以用统计学解释。前面我们已经说过，当购买者依靠自己的智力选择号码时，他们倾向于主动避免选择某些号码，结果是另外一些号码经常被选中。人为选择的结果使得各个号码被选中的频率非常不均匀。有一些号码几乎从来无人问津，而另一些号码非常受欢迎。计算机则完全不同。与人相比，计算机缺乏智慧、想象力和创造性，只会按照预定的结构高速、精确地运转。预定的结构——即计算机程序——忠实地产生几乎完全随机的结果，各个号码被选中的概率是没有差别的。这种随机性使得机器选择的稳定性远胜于人为选择。

这些统计结果为我们提供了宝贵的线索，帮助我们估计在大英彩票中哪些号码更受欢迎。假设我们想知道某一个特定的号码——比如说3——受欢迎的程度，我们可以挑出所有获奖号码中包括3的轮次，然后把这些轮次中的获奖人数累加在一起。用这个累加的和比较获奖人数的预期值（即就概率而言平均应有多少人获奖，这个数字可以反映一共售出了多少彩票)。如果累加和高于预期值，就说明这个号码比较流行；进一步说，累加和高出预期值的比例越大，则这个号码的流行程度越高。有几个研究人员利用这个思路（当然他们的具体方法可能更精巧）分析了大英彩票的中奖号码，得到了一些合理的结论。他们的结论与加拿大彩票的统计结论略有出入，但是保持了相当好的一致性。所以，如果你希望根据加拿大彩票的统计

结论估计大英彩票的情况，应当是可行的。然而，这种方法只能估计出具体的某个号码的流行程度，我们真正关心的却是各个号码组合的流行程度。

最引人注目的事件发生在大英彩票的第九轮，在那一轮竟有 133 人赢得头奖！如果根据那一轮售出的彩票数额估算，获奖人数应当是 5 人。假如彩票的购买者确实是随机地选择号码，这个结果足以证明统计学是一种错误的理论。所以，这个事实证明购买者的选择实际上不是随机的。为了解释为什么那么多人选择了这个获奖的号码组合，我们要观察一下大英彩票的外观，如图 2－2。

获奖的号码组合 {7, 17, 23, 32, 38, 42} 具备三个特征：

1. 各个号码出现在不同的行；

2. 每个号码都出现在居中的三列内；

3. 没有任何两个号码恰好相差 5，即没有两个号码在纵向相邻。

23

这些特征能说明什么？我的解释是：许多人在选择号码时就是拿着铅笔沿中间的三列从上往下划，并且略微左右移动，避免选中上下相邻的号码。他们认为，这样选择的结果就是随机的。（我在报纸上见过另外一种解释：多数获奖者出生于 1938 年和 1942 年。我认为我的解释比他高明得多。）

仅仅一个例子还不足以说明问题。其它的解释也是有可能的。然而，类似的模式出现过多次。为了仔细分析这个问题，我们必须把每一轮售出的彩票数额考虑进去。在一轮周末彩票中，通常售出 5 600 万至 7 000 万张彩票，在所有的 1 400 万中可能的号码组合中平均每一种卖出 4 至 5 张；在一轮非周末彩票中，通常平均每一个号码组合卖出 2 张。在我写这本书的时

图2-2 大英彩票的外观

1	2	3	4	5
6	7	8	9	10
11	12	13	14	15
16	17	18	19	20
21	22	23	24	25
26	27	28	29	30
31	32	33	34	35
36	37	38	39	40
41	42	43	44	45
46	47	48	49	

候，我发现至少还有三轮彩票的结果符合以上模式。有两轮周末彩票，获奖人数分别是57和20，另外一轮非周末彩票的获奖人数是16。这三个数比平均值（周末彩票的4至5和非周末彩票的2）高出好多倍。而除这4轮以外，没有任何一轮的头奖获得者超过20。当然，我的结论可能是错误的，不过如果你希望避免许多人与你瓜分奖金，我建议你不要用以上模式选择号码。

依据这种模式选择的号码趋向于均匀地分布在1到49之间。为了检验一组号码分布的均匀程度，我们可以计算一下间距最小的两个号码的差。在第九轮彩票的获奖号码中，间距最小的两个号码是38和42，它们的差是4。这个差我们称之为

24 “最小间距”，用英文字母M表示。当一组号码中存在相邻号码时，M=1。显而易见，M的取值范围在1和9之间，而且M的值越大，说明这组号码的分布越均匀。

在已知售出的彩票的总数的前提下，通过比较实际赢得头奖的人数与根据概率计算出的获奖人数的预期值，我们可以洞悉购买者选择号码的机理。对于每一轮，我们知道实际的获奖人数和预期的获奖人数。在此基础上，考虑每一轮的获奖号码的M值，把M值相等的各轮的结果加在一起，就可以得出一些统计学上的结论。

在最初的282轮抽彩中，共有132轮的获奖号码中包含相邻的号码，此时M=1。在这132轮中，共有330张彩票获得头奖。而就概率而言，赢得头奖的人数的预期值是514人。这说明，当获奖号码中包含相邻号码时，获奖人数的实际值较低，其结果是每张彩票分得的奖金数额较高。

表 2.4：282 轮大英彩票中的 M 值的统计规律

M 值	对应的轮次总数	实际获得头奖的人数	获奖人数的预期值	每张获奖彩票获得的奖金份额
1	132	330	514	较高
2	83	312	339	较高
3	31	144	113	较低
4	22	259	84	较低
5 及 5 以上	14	146	52	较低

表 2.4 中统计了不同的 M 值对应的情况。我们需要仔细观察第三列和第四列，比较这两列数字的大小。表中数字显示，当号码的均匀程度上升时（即 M 值增加时），实际的获奖人数增加，每张彩票获得的奖金份额下降。不过我们需要警惕一个事实：这个表的最后两列对应的轮次很少，分别是 22 和 14。由于涉及到案例个数太少，这个结论的代表性可能不好，当案例个数增大时结论可能有变化。另外应当注意，虽然号码组合 {1，2，3，4，5，6} 的 M 值已达到最低，但我们可以肯定，假如这个号码组合中奖的话，奖金将极低。这说明，仅以M 值作为标准是远远不够的，我们需要更多的线索。

类似的矛盾出现在其它场合。比如，统计一下购买者是否经常选择较大的号码（当然我们先要约定多大的数算“较大”）？假如我们约定，从 41 到 49 算“较大的号码”（即大英彩票的最后两行号码），统计结果相当古怪。当获奖号码组合中不包含较大的号码时，实际的获奖人数低于预期值；当获奖号码组合中恰好包含一个较大的号码时，实际的获奖人数高于预期值。但是，当获奖号码组合中包含 3 个或更多较大号码

时，实际的获奖人数极少。看起来，当购买者选择号码时经常在最后两列选一个号码，但是很少选 3 个或更多。迄今为止，还没有出现过获奖号码组合中包含 5 个或 6 个较大号码的情况，所以我们只能根据已知的结论大致猜想，这种情况下的获奖人数将极少。可是，我们已经知道，号码组合 {1，2，3，4，5，6} 非常受欢迎，我们很容易得出一个对称的结论：号码组合 {44，45，46，47，48，49} 一定吸引了许许多多的购买者。假如这个号码组合中奖的话，头奖奖金将少得可怜。

不过将来人们选择号码的方式可能有所变化。

不止一张彩票

你买的彩票越多，你中头奖的机会就越多。在一定程度上说，在一轮彩票中买多张彩票就相当于在同一轮赛马中同时在几匹马上下注——你其实在与自己赌博。然而，在一轮彩票中一次买许多张彩票与在许多轮彩票中各买一张彩票没有实质的差别，除非你打算一下买几十万张。

下面我们来证明这个结论。对于一张彩票，我们用字母 p 表示它中头奖的概率。前文的表 2.1 已给出了 p 的值，$p = 1/13983816$。如果你在同一轮彩票中买 N 张不同的彩票，你中头奖的概率等于 Np。

如果你在不同的 N 轮彩票中的每一轮各买一张彩票，由于每张彩票不中头奖的概率是（$1-p$），所以这 N 张彩票无一中奖的概率是（$1-p$）的 N 次方，即（$1-p)^N$。这 N 张彩票要么无一中奖，要么至少有一张中奖，所以你至少中一次奖的概率是

$$1-(1-p)^N$$

这个数与前面得到的 Np 不同，但差别极小。具体地说，当 N 小于 50 万时，这两个数的差低于 2%；当 N 小于 1 000 时，这两个数的差低于 2.5 万分之一。 26

如果你在几轮彩票中总共买 N 张彩票，除非 N 非常巨大，否则这 N 张彩票是在同一轮中买的还是在几轮中买的无关紧要，在具体的某一轮中买几张也没有关系——无论你如何选择，你获头奖的概率是一定的。当然你要注意，不能在同一轮彩票中买两次同一个号码组合。

有些人认为在同一轮彩票中买两次同一个号码组合是个好主意。在 1997 年 9 月的一轮彩票中，有 5 张彩票中头奖，其中两张属于肯特酒馆工会，而且这一轮的奖金中含有上两轮的累加奖金。他们真是非常非常幸运！他们本来的计划是首先选中 8 个号码，然后用 28 英镑买 28 张彩票，囊括这 8 个号码中的任意 6 个所构成的号码组合（在 8 个号码中选 6 个，恰好有 28 种可能）。可是最终，受委托填写彩票的人犯了一个错误：他漏掉了一个号码组合，仅选了 27 个号码组合，其中一个被选了两次。真是太走运了，被重复选择的号码组合中了头奖！由于这个巧合，他们用 28 英镑中了两个头奖，12 个“五号码奖”还有 14 个“四号码奖”。这个错误完全可能构成灾难——设想一下，假如那个填写彩票的人漏掉了头奖号码组合，而把一个只选中 4 个号码的彩票填了两次，那将是什么结果。

5 份奖金中的每一份是 540 万英镑，他们获得了两份。另外，12 个“五号码奖”赢得 2.4 万英镑。假如这个有趣的意外没发生，他们按照原来的打算买了 28 个号码组合，那么中头奖的彩票总数将下降为 4，每张彩票分得的奖金份额为 680 万英镑，他们的收入也就是这么多。然而，如果没有其他人选

中头奖号码组合，那么他们在头奖号码组合上买一张彩票和买两张彩票的结果是完全一样的。如果其他人买了一张头奖彩票，而他们持有两张，则他们赢得的份额是三分之二；如果其他人买了一张头奖彩票，而他们持有一张，则他们赢得的份额是二分之一。你预先无法知道哪个号码组合会赢，也不可能知道有多少人会中头奖，所以你应当买不同的号码组合——很明显，两个赢得头奖的二分之一的机会比一个赢得头奖的三分之二的机会更有吸引力。

各大新闻机关热衷于介绍这个传奇般的幸运事件，而且反复强调“这种事发生的概率只有‘19.6 亿分之一’”。我很奇怪这个数是怎么算出来的。我的结论是这样：28 张不同的彩票中 1 次头奖的概率是1 400万分之 28；28 张彩票选 27 个号码组合，其中一个组合买两次，则中一次以上头奖的概率是 1 400万分之 27；28 张彩票选 27 个号码组合，其中一个组合买两次，中两次头奖的概率是 1 400 万分之 1。无论如何，无法得出“19.6 亿分之一”这个结论。我相信，这个结论是由两个错误造就的：错误一，他们误以为“既然一张彩票中头奖的概率是 1 400 万分之 1，则两张彩票都中头奖的概率是1 400万分之 1 的平方；错误二，一个简单的算术错误——把 1 400 万分之 1 的平方算成了 19.6 亿分之一。[①]

如果你准备在同一轮彩票中买几张彩票，你应当认真考虑是否让你选中的号码组合中包含相同的号码。假如你买的所有彩票均包含几个相同的号码，而且其中一张彩票中了奖，那么其它几张彩票中奖的可能性也要上升。你可以买一个轮盘赌的

---

① 译者注：英国人习惯把 1 400 万计为 14million，把 19.6 亿计为 196million。对于英国人来说，“14million × 14million = 196million”是一个容易发生的错误。

转盘，利用这个玩艺帮助你选择几个幸运号码，比如说 15 个，然后就在这 15 个幸运号码中选择你的彩票号码。这样做的好处是显而易见的：如果某一张彩票中奖，则其它彩票中奖的可能性也大大提高；当然，这种做法的坏处也很明显：你在这轮彩票中一无所获的概率也在增大。综合起来看，好处和坏处抵消了。总而言之，假如你购买的彩票张数是一定的，无论你如何选择号码，从概率上说你有多少张彩票中奖是不变的。假如你的彩票中没有完全重合的号码组合，那么你赢得头奖的机会也是不变的。如果你期望像肯特酒馆工会那样，中一笔大奖，附带若干份小奖，这个期望是非常非常遥远的。如果你想深入讨论这个问题，你必须相信概率论的结论。

号码果真是随机产生的吗?

以上的讨论基于一个假设：在抽彩时机器产生的号码是随机的，全部 1 400 万个号码组合中的每一个最终被选中的概率是相等的。前文我已经说过，我们有理由相信这个假设是成立的。不过我们还是需要证明——我们怎么知道这里没有陷阱呢？也许某些小球比其它球重，也许某些小球里面有磁铁，也许整个机器里安装了遥控装置，也许某些人在幕后操纵彩票结果。这些情况当然是有可能的，但是我要告诉你，想在抽彩中作弊是非常困难的，而且一旦有人作弊，利用数学方法很容易发现。

28

细心的读者可能会指出：所有人造的机械都不可能完全避免误差，所以“各个号码组合出现的概率完全相等”是不可能的，因此机器产生的号码组合不可能是完全随机的。从理论上

说这种观点是正确的，但是没有什么实际意义。我们不需要保证每种号码组合出现的概率完全相等，只要证明没有人可以利用概率上的差别获利就足够了。就好像在足球场上两支球队用抛硬币的方法分场地，虽然从理论上说抛硬币的结果并非绝对的随机事件，但是此结果对于当事人是不可预测的。没有人可以利用概率上的不均匀作弊，所以整个过程是公平的。附录4中有详细讨论。

号码组合的全部可能性有1 400万之多，而我们所掌握的实际抽彩结果只有几百个。我们获得的信息如此有限，无法得出完整的结论。因此，我们把精力集中于两个问题：

(1) 每个单独的号码出现的概率是否相等？
(2) 前一轮的结果对后一轮是否有影响？

在我们掌握更丰富的数据之后，就可以讨论一些更复杂的因素，比如“在每一轮中某一对号码出现的概率是否相等”。我们不必使用非常规范的统计技术（在附录中介绍了这些技术），就可以简单地做出判断。

假设我们已经掌握了D轮抽奖的结果，并以此为依据判断各个号码出现的概率是否相等。在每一轮有6个号码被选中，所以我们总共得到6D个号码。对于全部的49个号码而言，每个号码被选中的频率的平均数是6D/49。我们用字母A表示这个数，即A=6D/49。有些号码被选中的频率比平均数高，有些则低。就我们现在讨论的例子而言，在1998年9月以前，一共举行过282轮抽奖，每个号码出现的平均频率A=34.5。这个数不依赖于具体某个号码的频率，即使某些号码出现的频率是0而另一些号码出现的频率超过50，平均频率依

然是 34.5。我们要检查一下各个号码出现的实际频率与平均频率接近的程度。以下是两个粗略的检验。

29

首先计算出 A 的平方根，以此作为判断“容许误差”的标准。当 A = 34.5 时，A 的平方根为 5.87，把这个数缩小一点并取整，得到 5.5。5.5 即为容许误差。计算具体的 49 个号码的出现频率与平均频率的差值，然后把差值与容许误差比较，即可得出结论。

第一个检验是判断处于容许误差范围内的号码是否达到 2/3。在此，容许误差的范围是 34.5 ± 5.5，即 29 到 40。我们期望在全部的 49 个号码中有 32 个出现的频率在 29 到 40 之间。如果频率在 29 到 40 之间的号码不超过 25 个，则视之为一个危险信号。事实上，恰好有 32 个号码出现的频率在 29 到 40 之间。第一个检验通过了。第二个检验是判断有多少个号码的频率与平均频率的差额高于容许误差的二倍。我们希望这个比例不超过 5%。具体地说，我们检查一下有多少个号码出现的频率低于 24 或者高于 45，我们希望结果不超过 3。事实上，在 282 轮抽奖中，没有任何一个号码出现的频率低于 24，只有一个号码出现的频率高于 45。第二个检验也顺利通过了。我们已经不必担心“每个单独的号码出现的概率是否相等”这个问题了。

为了精确地表达这些频率的均匀程度，我们需要一种定量的方法标记这 49 个号码对应的频率与其平均值的接近程度。附录 4 介绍了具体的方法。我们用一个量——W——来表示这个接近程度，在数学上这个量被称为“拟合良度统计量”。对于公平的彩票而言，W 的平均值是 48。当 W 值低于 48 时我们可以认为各频率接近平均频率的程度已高于我们的要求；而当 W 值高于 48 很多时，说明某些号码的频率与平均频率相差甚

远。附录 4 中的表 A4.1 描述了多大的 W 值在统计学上是可接受的。一个真正公平的彩票的 W 值高于 65 的几率只有 5%，这个 W 值在统计学上经常被用作检验“拟合良度统计量”的标准。在大英彩票的 282 轮抽彩中，W 值为 52.4，完全在可接受的范围内。

附录 4 中的表适用于彩票的轮次很多的情况。如果你掌握了某个彩票的 200 至 300 轮的结果，并且希望通过这些结果判断彩票是否公平，你可以用附录 4 介绍的方法计算其 W 值，并把算得的结果放到表中比较。不过你要小心：你得到的结果是一个公平的彩票的 W 值与你的判断对象的 W 值相符的概率。一个公平的彩票还有 5%的几率得出一个高于 65 的 W 值，如果你据此指控其中有人作弊可是缺乏依据的！假设你把一个已举行过许多轮的彩票分成若干段，分别计算每段的 W 值，你会发现单纯的随机事件会造就很多较大的 W 值。下文我们还要详细讨论这个问题。

也许你已经注意到了：我从来没有说过“某个彩票是绝对公平的”。统计学永远不可能得出像几何学中的“勾股定理”那样绝对的结论。当我们发现一个彩票的 W 值出现在可接受的范围内时，统计学的结论并非“啊哈，这个彩票是公平的”，而是“到目前为止没有证据表明这个彩票是不公平的”。这与 30 法庭上的情况类似：法官从不要求陪审团判断被告是否无辜，陪审团的职责是判断现有的证据是否可以“没有疑问地”证明被告的罪行。如果答案是“否”，则法庭的判决是“无罪”，即使陪审团相信被告在 75%的概率上是有罪的——75%还是不够大。在我们检验彩票的频率的过程中，我们始终不能做出最后判决。只有一种例外：在某个阶段出现了一个非常大的 W 值，使得公众普遍对彩票的公正性失去了信心，这时一定会换

一种新的抽彩方法。

在许多场合，拟合良度统计量是威力巨大的数学工具。如果在抽彩过程中真的有人做手脚，在拟合良度统计量上很容易露出马脚。但是需要小心，拟合良度统计量只是一种简便的计算方法，它不能区分正常的频率上的变动和人为的频率上的变动，这一点可能恰好是需要甄别的。假如你在检验过程中发现有一个号码出现的频率比平均频率高两个百分点，这已经是一个足以引起重视的差异，可是你也不能忽略这样一个事实：在一个公平的彩票中尚有 300 分之一的机会出现这种情况。附录 4 讨论了一个例子：一枚可能被做过手脚的硬币抛出正面的频率只有 47%，我们还是不能轻易地断言这个差异不是自然的。请仔细推敲这个例子。

以上讨论的都是各个号码在频率上的均匀程度，然而仅仅在频率上均匀还不足以说明问题。假设有一台生成获奖号码的机器，这台机器严格地按照从 1 到 49 的顺序依次发出号码，可以肯定在频率统计上结果是绝对均匀的，可是我们无论如何都不能把这样的号码生成过程称为“随机事件”。频率上的均匀程度不是惟一需要考虑的因素，一个更重要的因素是结果的不可预测性。

许多人潜心研究中奖号码的规律，并且做了好多图表，试图发现利用已知的数据预测中奖号码的方法。有一个学派主张：在上一轮中刚刚产生的号码是比较“热”的号码，在下一轮中这些号码再次出现的概率较高。另一个学派则提出针锋相对的观点：那些在最近没有出现过的号码在下一轮中出现的概率较高，因为只有这样才能符合所谓的“平均法则”——长期检验的结果必须证明每个号码出现的频率趋于平均值。很遗憾有两个问题谁都无法回答：各个号码如何能记住自己过去的历

史，从而决定自己在下一轮中的行为？如果统计图表要求某些号码出现，各个号码可能以什么方法使自己的行为与预测相符？我们的常识是“过去的结果与下一轮的结果无关”，当然，常识可能是错误的。从数学出发应当如何分析这个问题呢？

不同的理论对你的行为做出了不同的建议。如果“热号码”理论是正确的，你每次都应选择上一轮的中奖号码组合，这样你选中的正确号码的数量会高于平均值；如果“平均法则”理论是正确的，你每次都应选择过去出现的频率较低的号码，这样你获奖的机会将增加；如果每次中奖号码的生成过程与历史上的结果无关，你就应当选择随机的号码。

31

我们可以利用平均数来比较以上两个学派的理论。当我们买进一张彩票时，就概率而言在我们选中的六个号码中正确号码的个数是 36/49（六个号码中的每一个恰好是正确号码的概率是 6/49，所以选中的正确号码个数的期望值是 $6/49 \times 6$，前文介绍过算法）。假设我们在大英彩票的最近 281 轮的每一轮都买一张彩票，而且每次都选上一轮的中奖号码组合。就概率而言我们选中的正确号码的个数应当是 $281 \times 36/49 = 206.4$，而实际上我们选中的正确号码的个数是多少呢？如果“热号码”理论是正确的，这个数应明显高于 206.4；如果“平均法则”理论是正确的，这个数应明显低于 206.4。实际情况是这个数等于 190。无需复杂的统计学标准就可以得出结论：没有证据支持两个学派中的任何一个。

以上分析中仅仅检查连续的两轮中是否存在相关关系。可能有人会因此提出反对：也许各轮的结果之间的相关关系只是表现在长期的统计结果中，在相邻的两轮中并不体现。为了检验这种观点，我们设计一个复杂一点的实验。首先定义一个概念：如果一个号码出现的频率明显高于平均频率，则称之为

"热号码"。精确地说，如果一个号码在以前的7轮中至少出现过3次，就称之为热号码。这样，某些轮次中不包含热号码，另外一些轮次中则包含多个热号码。多数情况下，一轮中包含1到3个热号码。

我们统计一下从第8轮到第282轮的数据，其中各轮包含的热号码的个数分别是2，1，3，2，……，3，2，总数是557个。我们把这557个号码与特定的某一轮的获奖号码组合相比较，看看这557个号码中有多少个出现在获奖号码组合中。我们可以得到两个值：号码实际出现的个数和依据概率计算出来的平均值。如果我们相信各轮结果之间没有相关关系，我们应当期望实际的个数与平均值大致相等；如果我们相信"热号码"理论，我们应当期望实际的个数明显高于平均值；如果我们相信"平均法则"理论，我们应当期望实际的个数明显低于平均值。检验的结果如何呢？根据概率计算出的平均值是5 576/49 = 68.2，而实际的个数是67。所以，我们倾向于相信各轮结果之间没有相关关系。事实上，只有当实际的个数低于52或者高于84时，检验的结果才能支持"平均法则"理论或者"热号码"理论的假设。

我们选择"在以前的7轮中至少出现过3次"作为热号码的判定标准，这个标准是相当随意的。在我选择这个标准的最初，我并不知道检验的结果支持哪一种理论假设。当然我们也可以选择其他的频率作为热号码的标准，比如在以前的6轮中至少出现过3次，在以前的10轮中至少出现过4次，在以前的5轮中至少出现过2次。每个标准都将导致有价值的结论。

在任何时候我们都拥有数量一定的一批数据。我们可以用许多方法来定义热号码，然后依据热号码检验给定的数据。如果每一次检验都支持同一个理论假设，我们就可以认为，这个

假设获得了显著的证据支持。在具体的检验中，我们应注意一个问题：多大的数据偏差算是明显而异常的？一个具体的例子：一枚硬币连续抛出了10次正面，这算不算异常的偏差？这个问题必须在一定的数据背景下回答。假设我们一共就抛了10次硬币，而10次都得到正面，我们应当对事件的随机性持怀疑态度；然而，假如我们连续抛了几千次硬币，而在其中的某一阶段连续得到10次正面，那就没什么好奇怪的了。事实上，假如抛硬币的次数足够多，不出现连续的10次正面才是不正常的！同样的道理，如果我们完全随机地尝试许多不同的热号码的判定标准，比如在以前的12轮中至少出现过4次，在以前的3轮中至少出现过2次，等等，偶尔我们会得到一个不支持我们的理论假设的结果。

假设你收集了大量的数据，并且发现了一个不支持我们的理论假设的结论。你是否可以据此断言抽彩结果不是随机的？反过来说，在什么情况下我们可以断言这是一个自然而正常的现象？合理的办法是继续观察，看看未来的抽彩结果是否符合相同的规律。如果仅仅通过有限的观察否定抽彩的随机性，其态度是不科学的。由于大英彩票可以产生如此丰富的数据，会出现许多貌似异常的结果。下面是三个例子。

例一：在大英彩票进行了一年左右的时候，一些号码已经出现了10次以上，而号码39仅仅出现了一次。有人声称，这足以证明大英彩票的结果不是随机的。统计学的结论并非如此。如果你指定一个特定的号码，这个号码连续许多轮——比如说50轮——不出现的概率是极低的，概率值在0.0015左右；可是，**存在一个号码**连续50轮不出现的概率则比这个值高很多，概率值是7%左右。当然，7%也是一个小数目，但

已经不是**不可思议地小**。[①] 进一步说，39只是在连续40轮左右没出现以后才被列为怀疑对象，真正合理的检验手续应当仅仅统计39在被列为怀疑对象以后出现的频率。（当然此后39又出现了。它的纪录是连续52轮不出现。随着彩票的进行，随机性使得39终于摆脱了出现频率最低的状态，后来随机性又使得39重新回到出现频率最低的状态。在282轮抽彩之后，37和39两个号码并列于频率排行榜的底座上。连续52轮不出现的纪录后来被15打破，这个号码曾在第100轮出现，下一次出现在遥远的第166轮。这个新纪录很容易被其它号码刷新。）

例二：有人发现了存在一个由3个号码构成的号码组合，这个号码组合在连续的37轮彩票中出现了8次。一些报纸报道了这个发现，一个读者据此断言：大英彩票的结果不是随机的。这个断言不妥。关键在于，这37轮的选择是人为指定的。选择最初一轮和最后一轮的过程实际上带有明显的目的性：尽量使被选定的区间内符合条件的对象占更高的比例。这种选择检验对象的方法就是有问题的。如果把所有的抽彩结果都作为统计对象，我们会发现，并不存在一个频率明显高于平均值的由3个号码构成的号码组合。 33

例三：1996年夏末，从彩票的第86轮到第95轮，号码44在连续的10轮中出现了8次（实际上其中有两次是作为“额外机会号码”出现的）。对这个例子的分析类似于例二。同样，

① 这是一个复杂的概率问题。我们考虑两个概率值：A表示“存在一个号码连续50轮不出现的概率”，B表示“某一个特定的号码连续50轮不出现的概率”，A是B的49倍。关键在于前者没有指定特定的号码，因而有49种选择的可能。专家也经常在这类问题上犯错误——译者注。

当我们把所有的抽彩结果都作为统计对象时，并没有哪个号码的频率明显高于平均值。把额外机会号码考虑进去也一样。

这三个例子粗看起来都是令人吃惊的，然而经不起仔细推敲。最重要的差别在于这些引起我们注意的号码（或号码组合）是事后发现的，而非事先预言的。[①] 如果这些情况是被事先预言的，那可真有点麻烦了。有很多“特殊”的情况应当已经发生，实际上尚未发生。在迄今为止的 282 轮中，尚未发生连续两轮中包含相同的四个（或更多）相同号码的情况，根据概率，这种情况应当每 1 000 轮出现一次。直到第 267 轮才第一次出现全部六个号码都是奇数的情况，而根据概率，这种情况应当每 79 轮出现一次。我们可以设计出许许多多类似的例子，首先计算以下理论预期的频率，然后用实际数据检验一下。有时实际数据会偏高许多，有时实际数据会偏低许多，但是在大多数情况下，实际数据与理论预期符合得很好。

假如在抽彩结果的统计数据中发现了某些异常的现象，先别太得意。你应当从发现规律的那一刻起重新开始观察，看看你发现的规律是否在未来的抽彩中依然成立。如果确实成立，那么我必须恭喜你：你已经发现了抽彩活动的漏洞，你可以利用你发现的规律选择号码组合，从而提高你中奖的机会，至少你的收益会高于平均值。

---

① 概率论中有专门的概念讨论这种差别，即所谓的“验后概率”。有兴趣的读者可参考概率论方面的教科书——译者注。

小 结

大英彩票与其他国家的六合彩非常相似。组织者从彩票的购买者手中收钱，然后把其中的一部分作为奖金返还。对于购买者而言，回报远低于投入。在大英彩票中，回报率是45%。34 但是，价值一英镑的彩票可以带来数种潜在的中奖机会，抽彩活动已经造就了数以百计的百万富翁。此外，握着彩票而期待的过程可以给人快乐，这也是有价值的。

表2.2和表2.3给出了各个档次的奖金所对应的中奖机会和奖金额度。98%的彩票一无所获，95%的奖金额支付给末奖（10英镑）。只有头奖（1 400万分之一的机会）和额外机会奖（230万分之一的机会）的奖金额度达到了足以改变人生命运的程度。有些号码组合经常被购买者选中，如果一旦这类流行号码组合中奖，头奖奖金会令人失望地少。研究大英彩票和其他国家的彩票中的统计数据可以帮助你避免选择流行号码组合，这样你就不会面对数额可怜的头奖奖金。流行号码组合不是固定不变的。为了真正避免流行号码组合，最好利用器械帮助你选择完全随机的号码。你利用自己的大脑产生的“随机号码”并不是真正随机的，很可能别人也利用了相同的选择模式，并且选择了同样的号码组合，这样你选择的号码组合很可能就是流行号码组合。在存在累计奖金或奖金份额增加时你的回报率会高于45%，但是引起改变的仅仅是头奖奖金，其它奖金的中奖机会并不增加。所以除非你选中全部的6个号码，奖金额度的增加不会带给你什么具体的好处。如果你把一笔数额固定的钱投入彩票中，那么全投入到同一轮彩票中或分批投

人几轮彩票中没有明显的差别。

没有明显的证据表明某些特定的号码或号码组合出现的频率比平均值高（或低）。反过来说，假如真的存在这样的号码

35 或号码组合，通过统计方法很容易发现。

第三章　足球彩票与增值彩票

足球彩票

1923 年，三个曼彻斯特的电报员开始组织第一次足球彩票。由于担心被他们的老板发现，他们隐藏了身份，用了一个假名字——利特伍得（英文的意思是“小森林”）。他们在列队等候观看曼彻斯特联队比赛的球迷中发放了 4 000 张彩票，第一周只吸引到 35 个下注者。几周以后，1 万张彩票中的一张中了奖。第一个赛季结束的时候，发起者发现他们的事业亏本了，两个人决定退出，可是约翰·莫里斯决定坚持下去。开局不利，但是足球彩票最终获得成功。现在，有 1 000 万人定期购买足球彩票。虽然有些人模仿约翰·莫里斯设立了类似的彩票，利特伍得彩票依然在这个领域占据主导地位，它的奖金额也是最高的。

到 1928 年时，利特伍得彩票的每周奖金额已达到 1 万英镑。由于购买者不断增加，加上通货膨胀的影响，头奖的金额

一涨再涨，1950 年达到 6 位数。那时候，足球明星的工资不过是每周 10 英镑，披头士乐队的成员还穿着短裤。一个普通人想要发家致富，首先就会想到足球彩票。如果你在某个周六猜对了仅有的八场打成平局的比赛，你就可以用一便士换来终生受用不尽的财富。当然，许多人都没有意识到这个机会是多么微小。

为了保持彩票的吸引力，从 1951 年起澳大利亚足球比赛也被作为下注的对象。1963 年，因为阿克泰克赛季的许多比赛被取消，还特别成立了一个委员会。这个委员会的任务就是为那些被取消的比赛编造对阵情况，从而使彩票可以延续。36 1972 年，第一笔达到 50 万英镑的奖金诞生。7 年以后，一个失业的理发师用 45 便士赢得了 75 万英镑的奖金——这是她第一次买足球彩票。1984 年，一群护士央求她们的病人帮她们选择了一组号码，这一组号码赢得 100 万英镑。1991 年，一个人用 54 便士赢回了 200 万英镑。大英彩票开始以后，足球彩票受到冲击，但是依然保持着巨大的吸引力。

历史上曾有过许多种抽彩活动，多数已被人遗忘。今天，利特伍得彩票中最有吸引力的是“三倍机会彩票”。三倍机会彩票的规则很复杂，购买者需要在 49 场给定的足球赛中选择 8 场，尽量选择那些打成平局并且进球数不为 0 的比赛。这种彩票始于 1946 年，其名称的由来是因为它的奖金分三次评出。在一张彩票中，每选中一场打成平局并且进球数不为 0 的比赛，彩票计 3 分；每选中一场打成平局并且进球数为 0 的比赛，彩票计 2 分；每选中一场分出胜负的比赛，彩票计 1 分。为了增加趣味性，每张彩票的总分计算两次：一次根据终场得分情况计算（这是产生头奖的依据），另一次根据半场得分情况计算。

彩票的奖金中要扣除收入所得税，另外，彩票的组织者的花费和利润以及比赛监督机构的酬金也要从奖金中支出。每一周彩票的奖金占彩票的销售额的比例各不相同，但通常低于25%。一个典型的例子：在某一周售出了400万张彩票，总销售额在800万英镑到900万英镑之间，奖金总额为200万英镑。组织者首先在奖金中抽出一部分作为第一批奖金（半场成绩奖金），这笔钱大约占奖金总额的八分之一，其评奖依据是半场得分情况。如果在这一次评奖中没有人得满分，这笔奖金累加到第二批奖金和第三批奖金中。第二批奖金占总数的四分之三，第三批奖金占总数的四分之一，这两批奖金的评奖依据是终场得分情况。其中第三批奖金又被分成三个档次。全部奖金本周发放，不累加到下一轮。表3.1是一个直观的例子，记录的是1997年7月的三周的结果。

表3.1

	7月12日	7月19日	7月26日
一等奖	1人获得966 000英镑	263人各得3 700英镑	5人各得205 000英镑
二等奖	6人各得14 500英镑	3 165人各得31英镑	324人各得490英镑
三等奖	207人各得400英镑	30 480人各得2.6英镑	994人各得147英镑
四等奖	975人各得145英镑	155 381人各得1英镑	22 701人各得9英镑
半场成绩奖金	15人各得16 000英镑	8人各得9 000英镑	无人获奖

37

这三周的彩票依据澳大利亚足球比赛的结果抽奖，它们吸引的赌注比冬天时多。6月12日，获奖者仅仅1 200人；而下一周，获奖者接近20万！获奖人数的变化如此之大，引起了奖金额度的巨大变化。如果你在6月19日获奖，即使你赢得头奖，奖金额也少得不足以令你激动。你应当注意这种情况：

在某一周头奖奖金接近 100 万英镑，可是下一周的头奖奖金少得刚刚够你装潢厨房，或者供你的女儿读一年大学。在表 3.3 中，有一周的二等奖和两周的半场成绩奖已经超过了 7 月 19 日的头奖。有时候，由于非常多的人获得二三四等奖，使得他们的奖金额低于 1 英镑，此时将重新调整奖金分配方案，提高奖金额。

一等奖的奖金额度很大程度上依赖于打成平局的比赛的局数。有些购买者非常细心地把他们的足球知识运用到选择号码的过程中，仔细评价每支球队的实力，尽量选择那些参赛双方实力相当的比赛。当然，绝大多数人的选择没这么细致。实际上，如此煞费苦心地选择号码所获得的成效是相当有限的。80%以上的人多年来始终坚持选择固定的号码，或者每次完全随机地选择号码。迄今为止，最高获奖纪录由一个女士保持。你知道她是如何获得 200 万英镑的巨款的吗？18 年以来，她始终坚持买一组固定的号码，而且从未赢得任何奖金，直到最后一次的成功。如果你选择的号码集中于 1 到 31 之间，许多人与你分享奖金的可能性会明显上升。造成这种情况的原因有二：其一，很多人用出生日期作为幸运号码，这使得 1 到 31 之间的号码被选中的频率极高；其二，31 以上的号码对应的比赛场次往往是英国国内的弱队或苏格兰球队，而买彩票的人主要集中于英国，他们通常宁愿选择自己比较熟悉的国内强队的比赛。总的说来，多数人随机地在 49 个号码中选择 8 个，中奖与否，听天由命。

38 在 49 个号码中选择 8 个，全部的可能性计为${}^{49}C_8$，这个数略大于 4.5 亿（参阅附录 1）。由于每张彩票的售价仅仅是三分之四便士，购买者通常一次买进好多张彩票（平均花费 2 英镑）。最常见的方法是一次挑选出 10 个号码，在 10 个号码中

选择 8 个共有 45 种可能（即$^{10}C_8$），购买者一次买进 45 张彩票，覆盖全部 45 种可能。此时购买者要付出 60 便士，根据彩票规则，这是购买彩票时的最低消费标准（每次至少买进 45 张彩票）。80%以上的购买者利用这种方法选择彩票：首先选择一组号码（超过 8 个），然后买进这组号码的全部可能组合。这种购买方法使得彩票的组织者很容易统计获奖情况。如果你选择的这组号码中包括 8 个最佳号码（即实际上打成平局并且进球数不为 0 的比赛的场次），毫无疑问你已经得了最高分；如果选择的这组号码中包含的最佳号码个数小于 8，计算你的得分也很容易。例如，你选择的 10 个号码中包括 7 场打成平局并且进球数不为 0 的比赛，2 场打成平局并且进球数为 0 的比赛以及 1 场分出胜负的比赛，则你的 45 张彩票中有 2 张得 23 分，8 张得 22 分，14 张得 21 分，21 张得 20 分。

表 3.2

你选择的号码个数	9	10	11	12	13	14	15
可以构成的八号码组合的个数	9	45	165	495	1287	3003	6435
总金额	0.12	0.60	2.20	6.60	17.16	40.04	85.80

表 3.2 可以帮助你计算你应当为你的选择付多少钱。第二行表示你选择的那些号码可以构成多少个八号码组合，第三行是以每张彩票价值三分之四便士计算出的总金额。如果你愿意，你可以同时买进某一个八号码组合很多次，一旦这个号码组合中了头奖，你所分得的奖金份额会增加；不过如果你是惟一的获奖者，这样做不会给你带来任何好处，你的奖金不会变得更多。

除了覆盖所有八号码组合以外，最流行的方法是首先选择若干个号码，然后在其中选择一些八号码组合（但不覆盖所有八号码组合）。至于在号码已经圈定以后，如何选择八号码组合，彩票的组织者和一些报纸提供了许多详尽的方案。这类方案通常称为“标准方案”。例如，彩票公司提供一种选择方案——利特方案，这种方案要求你圈定 16 个号码，然后它会帮助你用这 16 个号码组成 60 张彩票。假如你想应用这种方案，你只需在你的彩票上圈定 16 个号码，然后在彩票上写上“按利特方案选择 60 张彩票，共 80 便士”。彩票公司的计算机会自动记录你所选定的号码组合，开奖时还会自动统计你的每张彩票的得分。标准方案有很多种，无论你最初圈定多少个号码，都可以找到一种可应用的标准方案。在由报纸提供的标准方案中，最受欢迎的一种包含 600 张彩票，价值 8 英镑，要求你圈定 17 个号码。如果你想覆盖 17 个号码组成的所有八号码组合，你的花费将是 300 英镑。这种标准方案是通过周密的计算产生的，它可以保证：如果你圈定的 17 个号码中包含 8 个最佳号码，它所选择的 600 张彩票中至少有一张包括 7 个最佳号码（当然也不排除包括 8 个最佳号码的可能）。设计这样一种方案需要精妙的数学方法，在科学计算和商业统计中经常用到同样的方法。 39

其实标准方案与随机选择有些类似，但是看起来更合理。标准方案的第一个好处是可以帮助你节省劳动。假如你既不使用标准方案，也不选择覆盖若干个号码中的所有八号码组合，你就必须在彩票上圈出你想买的所有号码组合。如果你想买两英镑的彩票，你就必须圈出 150 个号码组合——那可是 1 200 个号码！使用标准方案可以减少许多麻烦；标准方案的第二个好处是可以使你的选择更有效率。在数学上标准方案是合理

的，至少它可以使你避免一种遗憾：也许你圈定的号码中已经包含足够多的最佳号码，可是由于你的选择方案不合理，使得这些最佳号码分散在不同彩票中，每张彩票的得分都不高；标准方案的最后一个好处在于它是免费的。

我们没法断定一种选择方案一定比另一种选择方案好。标准方案只是提供一种便捷地生成彩票的方法，至于你可以期望得多少奖金，仅仅依赖于两个因素：你预测比赛结果的能力，以及你买多少张彩票。无论你使用标准方案，还是选择覆盖若干个号码中的所有八号码组合，你最终选择的彩票中都包含很多交叉、重复的号码，这类选择的好处是一旦其中一张彩票中奖，其它彩票中奖的机会也相应增加——这与六合彩中的情况相似。那些令人羡慕的大赢家总是同时赢得一项大奖和许多项小奖。

在彩票中最吸引人的是头奖——头奖的意思是一个人独得一等奖。平均每个月产生一个头奖，但是头奖的产生是一个随机事件。历史上出现过这样的情况：有一个阶段连续 16 个月没出现头奖，而在另一个阶段，连续 3 周每周都有头奖。头奖大大地鼓舞下注者的士气，并且吸引更多的钱投入彩票，因此彩票公司总是想方设法增加头奖产生的概率。彩票公司可以通过调整每张彩票的售价影响头奖产生的概率，他们可以估算出大致有多少钱将投入彩票，然后通过调整彩票单价即可以控制彩票的销售量。销售量过高或过低都不利于鼓舞下注者的士气。因此，彩票公司在确定彩票单价时必须在一个复杂的均衡中做出选择。（这可以解释为什么最大的彩票公司的彩票定价是三分之四便士，而某些小公司选择十一分之五便士或八分之一便士。）

几十年以来，足球彩票的规则变化很大。50 年前，可供

下注者选择的比赛场次在55左右，评分系统是平局3分（不区分有进球平局和无进球平局）、客场胜2分、主场胜1分，要求下注者选出8场比赛。备选场次和评分系统与从前不同，可是彩票中有两样始终不变：其一，下注者需选出8场比赛；40 其二，彩票的最高得分是24分。历史上，备选场次曾多达59场，评分系统中也曾区分0比0平局、1比1平局、更高比分平局。现在，备选场次为49场，评分档次分为3分、2分、1分，有进球平局得分最高，无进球平局次之，主场胜和客场胜不加区分，得分最低。

彩票公司把备选场次定为49场，也就决定了彩票号码的全部可能组合一共有多少种（$^{49}C_8$）。事实上，这个数已经大致决定了将有多少张彩票售出。平均而言，每一个可能的号码组合售出一次。（这是一个粗略的说法。通常，每一个可能的号码组合售出的次数在1/3到3之间。）

彩票公司希望获得最高分的彩票数较少（因为这样每一份一等奖的奖金额会较大）。每一周备选场次中出现的平局次数各不相同，通常有1/3到1/4的比赛打成平局，而在平局中通常有1/4强的比赛没有进球。平均而言，每周有9到10场比赛打成有进球平局，4场左右的比赛打成无进球平局。在此条件下，出现头奖的概率是多少？

当有进球平局的局数不低于9时，获得24分的彩票将不低于9种。（利用表3.2可以得到精确的对应关系。表3.2的第一行可以当做有进球平局的局数，第二行可以当做获得24分的彩票的种数。）这种情况下出现头奖（即由一个人独得一等奖）的概率较低。有进球平局的局数为8时，只有一个号码组合可以得24分。此时，如果某一个下注者恰好选中了全部8个号码，则出现头奖的概率大大上升。然而，如果所有彩票

都没能得24分，则出现头奖的概率依然较低。考虑这个例子：某一周，出现8局有进球平局和一局无进球平局，并且没有人选中全部的8局有进球平局。此时最高得分为23分，获得23分的彩票可能性有7种（在8局有进球平局中选7局，再加上一局无进球平局），仅有一张彩票获得最高分的概率下降（降低到原来的七分之一）。如果无进球平局的局数大于1，则出现头奖的概率进一步下降。当有进球平局仅有7局时，出现头奖的概率较高。此时，获得23分的彩票的种数恰好等于这一轮中的无进球平局的局数。

我不再深入讨论平局局数变化时的各种情况。除了平局局数以外，平局所对应的号码也会影响头奖出现的概率，因为不同的号码被下注者选中的频率不同。（前文说过，号码不超过31时，被选中的频率较高，相应地出现头奖的可能性下降。）如果你想改善自己获得头奖的概率，需要透彻的分析。这不是件容易的事。 41

增值彩票

1956年，哈罗得·马克米兰向民众介绍增值彩票时说："增值彩票不是赌博，因为任何人都不会输。"在丘吉尔当政的时代，政府称这种彩票可以使人"无所失而有所得"，反对党却指责其为"肮脏的抽彩"。每张增值彩票的面值是1英镑，对奖之后会产生大量奖金，为了支付奖金，政府发行了很多钞票。增值彩票的第一轮始于1957年，当时的奖金不高，头奖仅为1 000英镑，但此后头奖奖金一涨再涨，1971年达到5万英镑，1980年达到25万英镑，1994年以后达到100万英镑。

中奖的机会很多，只要你持有一张彩票，你可以常年拥有中奖机会。

这些年以来增值彩票的规则经常变化，通常由民意测验决定。购买增值彩票的投资回报率与国家银行的储蓄利率大致相等。1998年，国家银行的利率是每年5%，而投资彩票的收益期望值也是这个数。与其它彩票相比，增值彩票的独特之处在于它本身是有价的，在任何时候，你都可以把彩票按原始面值兑换成现金。增值彩票的另一个好处是一旦中奖，你不必支付所得税。如果不中奖，你也没什么损失——如果你不在乎存款利息。你每买一张彩票，从下个月开始你就可以参加对奖，但是每个月只能中奖一次。在最初的许多年里，一次可以只买一张彩票，但是后来，为了提高抽彩的效率，政府规定每次购买至少要买100张。许多父母（或祖父母）把增值彩票作为圣诞礼物或生日礼物送给孩子。

根据1997年7月的统计数据，每年购买增值彩票的投资回报率为4.75%，产生的奖金总额大约84亿英镑，平均每个英国成年人获200英镑。当月产生的奖金额为3 300万英镑，分为439 875份奖金。奖金额度及对应的获奖人数如下：

1人获100万英镑
5人获10万英镑
8人获5万英镑
19人获2.5万英镑
47人获1万英镑
93人获5 000英镑
1984人获1 000英镑
5 952人获500英镑

64 479 人获 100 英镑
367 287 人获 50 英镑

每个月固定地产生一个 100 万英镑的头奖，这个数字保持着巨大的吸引力。如果有 10 万个人，每人持有 2 万张彩票（根据规则每个人持有的彩票数不得超过 2 万），而其中的每个人每个月赢得头奖的概率是 40 万分之一。有四分之三的奖金分配给等级最低的两档奖金，这两档奖金的获奖人数占全部获奖人数的 98%左右。每个月单独的一张彩票赢得 1 000 英镑以上的奖金的概率大约是 400 万分之一。这个概率不够大，不过你不必因此沮丧，因为你每个月都有获奖机会。平均而言，每张彩票每年带给你的收益略低于 4 便士。

1997 年 7 月赢得头奖的彩票的编号是 69GK540100。每张彩票的编号包括十位数字或字母，当首位数字为 0 时通常把 0 省略。正如一个人的身份证号码中包含了一个人出生日期的信息，彩票的编号中也包含了一些彩票本身的信息。在编号 69GK540100 中，第三位字母 G 表示这张彩票是 2 000 张同时售出的彩票中的一张。第四位的位置上只能出现十个字母中的一个，这十个可能的字母是 Z，B，F，K，L，N，P，S，T 以及 W，其实我们完全可以用从 0 到 9 的十个数字替换这十个字母。所以，这个编号表示面值为 2 000 英镑且编号为 693540100 的彩票。

获奖号码由一种被称为“电子随机号码发生器”的仪器生成。第一台仪器在 1973 年退役，现在使用的机器是第三台，于 1998 年开始服役。严格地说，绝大多数所谓的随机号码发生器生成的号码并不是真正的随机数，只是看起来像是随机的。这些号码可以通过随机性检验，看起来是完全随机的，但

是程序员可以控制号码的顺序，并可以预测将要产生的号码。利用一台这样的仪器来决定中奖号码显然是不可行的，幸好增值彩票中使用的电子随机号码发生器并不是这样的仪器。

电子随机号码发生器是根据噪声二极管中的自由电子运动设计的，而自由电子的运动是完全随机和不可预测的。通过比较两个噪声二极管中的噪声强度可以产生一系列的脉冲信号，在一段确定的时间内发出的脉冲信号的个数可以精确统计，这个数通常在1万至1.5万之间。这个数的最后一位数字被取出来用于产生中奖号码，两个这样的数字决定最终号码中的一位。

在电子随机号码发生器生成一个中奖号码之后，另外一台 43 计算机负责判断此号码是否符合中奖条件。以下三种号码不符合中奖条件：此号码对应的彩票尚未售出；此号码对应的彩票已售出但被兑换为现金；本月此号码对应的彩票已经中过奖。机器产生的第一个符合中奖条件的号码获得100万英镑，随后产生的号码依次获得数额较小的奖金，直到全部奖金分配完毕。购买者普遍担心的一个问题是“电子随机号码发生器是否会漏掉我的号码”？这个疑虑是完全多余的。抽彩过程由两台机器完成，第一台机器负责生成号码，第二台机器负责判断号码是否符合条件。电子随机号码发生器只是在全部可能的号码中随机地选择号码，它生成的号码既包括符合中奖条件的号码，也包括不符合中奖条件的号码。漏掉某个号码的可能性并不存在。

从理论上说这台机器无懈可击。另外，国家统计局进行过严格的随机性检验。检验结果已公开发表，其中包括三个结论：

1. 从 0 到 9 的各个数字出现的频率相等；

2. 对于每一个确定的数字，其下一位出现哪一个数字是随机的；

3. 取出中奖号码中的连续四个号码组成一个四号码组合，这个四号码组合出现的频率符合平均频率。类似地，三号码组合、两对号码组合、两号码组合和十号码组合出现的频率同样符合平均频率。

希望了解技术细节的读者可以参考附录 4。现在最后一个需要检验的问题是仪器产生的号码的顺序是否存在固定倾向。这个问题非常重要，因为第一个产生的号码对应的奖金远远高于其它号码。检验这个问题必须考虑彩票编号的第三位，因为第三位的值反映特殊的意义。在实际检验中所有号码根据第三位的不同分组，考虑了每个组售出的彩票的数量。每个月抽彩开始以后，国家统计局都要分析中奖号码的顺序是否存在固定倾向。迄今为止结论是否定的。

虽然经过了如此严格的检验，还是有许多人对抽彩的公正性心怀狐疑。一个常见的质疑是多数获得大额奖金的人来自南方。这是事实，但是不能说明任何问题，因为南方的彩票销售额高于北方。另一个质疑是许多获奖号码具备同一个特征：获奖号码的最后六个数字中有一对相邻的数字，其数字完全相同。比如，在最初的 40 轮彩票中就有 16 轮具备这个特征。粗看起来有点奇怪，仔细一想就会发现这不足为奇。在最后六个数字中可以组成五对相邻数字，每一对数字有 100 种可能（从 00 到 99），其中有 10 种由相同的数字组成（00，11，22，33，44，55，66，77，88，99），另外 90 种由不同的数字组成。所以每一对相邻数字由不同数字组成的概率是 9/10。这五对相

邻数字都由不同数字组成的概率是 $(9/10)^5$，大致等于59%，也就是说其中至少有一对相邻数字由相同数字组成的概率是41%。16/40不是和41%非常接近吗？还有一个著名的反对意见：中奖号码极少以0开头。这是事实，但是我们必须注意另一个事实：许多以0开头的彩票在增值彩票创办的初期发行，其中很多已经兑换成现金。抽彩规则规定彩票不可以作为遗产继承，在彩票的主人去世以后彩票必须兑换为现金。

你获奖的概率有多大？你赢得的奖金的期望值是多少？简单地说，你持有的彩票越多，你的机会就越好。但是你必须面对一个事实：即使你持有2万张彩票（这是规则规定的极限值），非常背运的可能性也是存在的。据统计，1996年有2 000人持有2万张彩票，平均每人中奖13.6次。其中一个幸运儿中奖达27次，另外一个倒霉蛋一次奖也没中。

目前购买增值彩票的投资回报率是每年5%，而银行的存款利率是每年8%。存款利息要上交收入所得税，而彩票收入是免税的，扣除所得税以后这两个值相等，差别在于彩票的收入是不可预期的。很多人宁愿把富余的钱投入彩票。在很多经济宽裕的家庭中，夫妇二人都持有2万张彩票。如果一个人持有1 000张彩票，他每年的彩票收益平均为50英镑。在每100个这样的人之中，一年中平均53人一无所获，34人中奖一次（通常是50英镑），11人中奖两次，2人中奖三次或更多。

假设在1997年7月时有10万人持有极限额的彩票并已保持一年，则其中每个人获奖次数的期望值是12.6，概率最高的情况是获奖12次。但是，其中的四到五人将在整整一年以内一无所获，而另外的四到五人将获奖29次以上。

你在今年的收入对明年没有任何影响。电子随机号码发生器不关心你曾经得过多少奖金。如果你希望获得稳定的投资回

报，增值彩票不是理想选择；然而，如果你期待意外的惊喜，不妨试试。每个月都会造就一个百万富翁，谁敢肯定那不是你呢?

最佳选择?

在足球彩票和增值彩票之后，大英帝国彩票提供了第三个一夜暴富的机会。除了巨额奖金的诱惑以外，三种彩票还保持着独特的吸引力。狂热的球迷热衷于足球彩票，因为这种彩票与他们本人的爱好相符，而且他们的奖金的一部分将捐助给足球事业，这使他们很高兴。大英帝国彩票所吸引的资金的28%将捐助给慈善机关，这使得大英帝国彩票的购买者有一种道义上的满足感。增值彩票的持有者可以认为自己为国家做出了贡献——使储蓄额增加。这三种彩票对购买者的隐私权都很尊重。 45

彩票的组织者为方便购买者做了大量努力。足球彩票拥有一个由7万人组成的网络，负责彩票和奖金的流通。增值彩票直接在邮局发行，任何时间、任何邮局都提供服务。大英帝国彩票在全国设立2万个零售点，购买者也可以购买季票，即一次性地同时买人未来许多轮的彩票。

这三种彩票对购买者的技术要求不同。增值彩票是最简单的，不需要任何技巧。买增值彩票就像在超市里买洗涤液一样，随便挑一个就行了，反正你的运气是由机器随机决定的。大英帝国彩票要复杂一些，因为你要尽量避免选中“流行号码组合”。足球彩票需要的技巧更多，如果你对足球的了解非常多，你的知识对于选择彩票无疑是有帮助的。

如果你在大英帝国彩票中获奖，你必须在180天以内做出声明，并且提供有效收据。没有人追着你给你送奖金。许多奖金因为无人认领而作废，其中包括一笔高达200万英镑的头奖。每周大约有100万英镑的奖金作废。彩票公司保留着所有彩票的存根，所以他们完全可以找到获奖者，但是如果你没有在规定期限内领奖，那是你自己的责任。增值彩票则不同，你根本不必检查自己是否在获奖名单上，老老实实地在家等着就行了。当然，如果你的地址有变，最好通知有关部门。自1957年以来，有25万份奖金因所有者地址变更无法发放，总额达1 500万英镑。有两份价值25英镑的奖金在抽奖的第一轮产生，至今仍等着获奖者领取。增值彩票没有领奖期限的限制，不过如果你在18个月以内没有领奖，国家储备局将不再接受你的查询。你可以委托私立的商业机关查询，当然，要缴纳查询费。当你选择购买哪一种彩票时，这些因素都是需要考虑的，但是最重要的因素只有两个：平均的投资回报是多少，以及你中奖的机会有多大。

首先我们比较一下投资回报率。假设你在三种彩票中分别投入100英镑，一周以后你的平均回报分别是：足球彩票25英镑，大英帝国彩票45英镑，增值彩票100.10英镑。优劣是很明显的。

增值彩票的性质与另外两种彩票不同。为了使比较更有说服力，我们对比3种情况：情况一，投资5 000英镑买增值彩票；情况二，把5 000英镑存入银行，每周取出利息买足球彩票；情况三，把5 000英镑存入银行，每周取出利息买大英帝国彩票。在第一种情况下，你每年的利润率是5%。平均每一年你将获得三笔50英镑的奖金，再加上获得高额奖金的机会（当然这种机会很小），你每年的收益是250英镑。

在第二种情况下，把5 000英镑存入银行，每周的利息是5英镑。每周在足球彩票中投入5英镑，按照25%的回报率，你每年的收益是65英镑。在第三种情况下，每周在大英帝国彩票中投入5英镑，按照45%的回报率，你每年的收益是117英镑。在大英帝国彩票中你可以应用某些技巧使收益提高。比如说，你可以在头奖奖金额较高的轮次中多买一些彩票，而在其它轮次中少买一些彩票。不过即便如此，你也只能把回报率提高到60%，你每年的收益是156英镑。其实无论投资额是多少，三种彩票的优劣次序是固定的：增值彩票最佳，大英帝国彩票次之，足球彩票最差。

此外我们还应当比较一下在三种彩票中赢得巨额奖金的机会。事实上，多数人在买彩票时并不关心平均回报是多少。当然，赢得巨额奖金的机会是微小的，但是多数人真正在乎的恰恰是这个微小的机会。因此，比较这三种彩票给予我们的赢得巨额奖金的机会，其意义更甚于比较平均回报率。我们还是假设投资额为5 000英镑，在增值彩票中用这笔钱买彩票，在大英帝国彩票和足球彩票中用银行利息买彩票。我们想回答两个问题：第一，连续20年投资彩票，赢得100万英镑以上的奖金的概率是多少？第二，连续20年投资彩票，赢得10万英镑以上的奖金的概率是多少？下面的分析基于1997年末以前的数据，读者可以利用最新的数据结合下文介绍的方法更新我的结论。 47

在增值彩票中，连续20年持有5 000张彩票，每个月开奖一次，因此有120万次中奖机会。彩票的总数是84亿，每次开奖在其中产生一个100万英镑的大奖，所以赢得100万英镑以上的奖金的机会是120万比84亿，这个概率约等于1/7 000。在大英帝国彩票中，平均每2 250万张彩票中有一张获得100

万英镑以上的奖金。20 年之间，你的存款利息可以买 5 200 张彩票，赢得 100 万英镑以上的奖金的机会是 5 200 比 2 250 万，这个概率约等于 1/4 300。计算足球彩票中的情况相对复杂，因为近年来足球彩票的销售额逐年下降，而我们依据的是过去的数据，这使得我们的计算结果可能比实际情况乐观。为计算方便，近似地认为头奖奖金的平均值是 100 万英镑。每个月你可以买 1 600 张彩票，而足球彩票的月销售量为 25 亿。在 20 年之内，你赢得一次头奖的概率是 1/6 500。综上所述，如果你的目标是梦幻般的百万英镑，大英帝国彩票是最佳选择，足球彩票和增值彩票则相差无几。

假设你的目标略微现实一点，只是赢得 10 万英镑，计算方法是类似的。在增值彩票中，赢得一份 10 万英镑的奖金的概率是 100 万英镑的六倍左右，所以你实现目标的概率是 1/1 200。在大英帝国彩票中，400 万分之一的彩票可以赢得 10 万英镑以上的奖金，所以你的机会是1/800。在足球彩票中，10 万英镑的奖金出现的频率是 100 万英镑的六倍左右，所以你的机会是 1/1 100。

如果你的预期目标是 10 万英镑或 100 万英镑，投资额是 5 000英镑，选择大英帝国彩票是明智之举，选择增值彩票或足球彩票则机会小很多。当你的预期目标、投资额、银行利率、投资期限等因素变化时，结论会有所不同，但是计算方法是不变的。你可以根据自己的情况计算你的机会，基本原则是算出每次对奖中你实现目标的概率 p 和你参与对奖次数 N，如果 N 和 P 的乘积小于 0.1，你实现目标的概率大致等于这个乘积。其实你完全可以把你的资金存入银行，这是最现实的选择，当然，你得不到买彩票的乐趣。

第四章　一枚硬币，多种游戏

有时候一枚硬币仅仅是一枚硬币，有时候一枚硬币是一个理想的随机事件发生器。当你和太太为看什么电视节目或圣诞节到谁的父母家度假争执不下时，你们可以通过抛硬币决定——硬币的地位是纸币和信用卡无法取代的。你们可以约定：如果硬币出现正面，按你的决定办；如果硬币出现反面，按她的决定办。如果真的出现反面，你还有一条退路——三局两胜。抛硬币的每一种结果代表一个决定，这个结论未必是最终的决定，但是它确实能帮助你做出最后的决定。事实上，许多游戏的原理与抛硬币相同，这些游戏可以用同样的方法分析。

任何硬币都不是完全对称的，所以出现正面的概率和出现反面的概率很可能并不完全相等。不过不要紧，只要任何两次抛硬币的事件相互独立，我们就可以“设计”一个理想硬币。我们每次连续抛两次硬币，抛两次算作一轮。如果两次都抛出正面或都抛出反面，则这一轮结果视为无效，再来一轮；如果第一次是正面而第二次是反面，则这一轮的结果视为正面；如果第一次是反面而第二次是正面，则这一轮的结果视为反面。由于第一次是正面而第二次是反面的概率与第一次是反面而第

二次是正面的概率完全相等，所以这一轮的结论相当于一枚理想硬币。下文的讨论全部基于理想硬币。

错误的直觉

普通人关于随机事件的直觉往往是错误的。我们从一个简单的例子入手：连续抛六次硬币，得到的结果有 $2^6=64$ 种可能（参阅附录 1），分别是：

正正正正正正
正正正正正反
正正正正反正
……　……
反反反反反正
反反反反反反

49

由于硬币是理想硬币，所以每一种结果出现的概率完全相等。全部出现正面的概率是 1/64，全部出现正面或全部出现反面的概率是 2/64 = 1/32。现在我们想知道一个特定的概率：得到的正面的次数与得到的反面的次数一样多的概率是多少？我们先要算出在全部 64 种可能结果中有多少种正面恰好出现三次，这个数是 $^6C_3=20$。所以我们的答案是 20/64 = 5/16。

如果抛硬币的次数不是六，结论会有何变化？显然，只有抛硬币的次数是偶数时我们才可以问“得到正面的次数与得到反面的次数一样多的概率”。假设抛硬币的次数分别是 2，4，6，8，10，……，等等，随着抛硬币次数的增加，得到正面的次数与得到反面的次数一样多的概率是否递增？很多人的直觉是“是”。大错特错！正确的答案恰恰相反：随着抛硬币的次

数的增加，得到的正面的次数与得到的反面的次数一样多的概率递减。

计算方法很简单。假设抛硬币的次数是 2n，则全部的可能结果有 2^{2n}种，其中有^{2n}Cn 种符合条件，答案是$^{2n}Cn/2^{2n}$。表 4.1 给出了最初的 5 个 n 对应的结论。为了便于比较，统一各个分数的分母，列在第三行。显然，结论是递减的。

表 4.1：抛硬币的次数与对应的概率值

抛硬币的次数	2	4	6	8	10
对应的概率值	1/2	3/8	5/16	35/128	63/256
统一分母后的结果	128/256	96/256	80/256	70/256	63/256

这组数据内部存在数学规律。第二个数与第一个数的比值是 1/2，第二个数与第一个数的比值是 3/4，以此类推，从第三个数开始每个数与前一个数的比值分别是 5/6，7/8，9/10，等等。这个规律对于每一个 n 值都有效。这也可以说明，我们关心的概率值是递减的。

1730 年，詹姆斯·斯达灵发表了一个极其精妙的计算公式，即斯达灵公式。这个公式提供了一种非常有效的近似算法。假设我们抛 100 次硬币，得到 50 个正面和 50 个反面的概率是多少？根据斯达灵公式，这个概率非常接近于$1/\sqrt{(50\pi)}$，约等于 8%。

π 居然出现在这个公式里，看起来很奇怪，我们的问题好像跟圆周率扯不上任何关系。不过 π 是一个非常活跃的数学常数，很多场合都免不了它的出席。假设抛 20 次硬币，得到 10 个正面和 10 个反面的概率是多少？根据斯达灵公式，这个数

是接近于 $1/\sqrt{(10\pi)}$，约等于 0.178 4，而这个概率的精确值是 0.176 2，两个值符合得非常好。斯达灵公式的一般形式是：

抛 2n 次硬币，得到 n 次正面和 n 次反面的概率非常接近于 $1/\sqrt{(n\pi)}$。

当 n 的值较大时，这个近似公式的符合程度更佳，其计算结果总是比精确值略大一点。

利用斯达灵公式可以很方便地得出某些结论。比如，我们可以知道当 n 达到多大时，正面出现的次数与反面出现的次数一样多的概率。如果我们希望这个概率值低于 1/100，只要使 $1/\sqrt{(n\pi)}$ 小于 1/100 就行了。计算结果是需要抛 6 366.2 次，我们取一个比较“整”的数，并且调大一点，6 500 就很好地满足我们的要求。当我们需要一个估计值时，取整是很好的办法。

如果我们希望这个概率低于 1/1 000，只要抛 65 万次以上就可以了。其实 64 万也可以，不过为了有把握，我们宁愿取比较大的数。当抛硬币的次数高于 100 万时，正面出现的次数与反面的次数相等的概率已经非常非常低了。但是，正面出现的次数非常非常接近于 50%。

硬币游戏

最简单的硬币游戏可能是这样：甲和乙各出 1 英镑，而后某个人抛硬币。如果结果为正面，2 英镑归甲；如果结果为反面，2 英镑归乙。这个游戏是公平的，但相当枯燥。怎样使游戏变得更有趣呢？

51

圣彼得堡游戏是一个有趣的变种，其规则如下：你的对手

先付给你一笔钱，然后你抛硬币，直到抛出一次正面为止。记住你用了多少次才抛出正面，这个次数决定了你要付多少钱。有时候你只用一或两次就能得到正面，有时候需要很多次，连抛 20 多次才得到正面的情况也非罕见。你用的次数越多，你要付的钱也越多。假如你一次就抛出正面，你付给对手 2 英镑；假如你两次才抛出正面，你付给对手 4 英镑；假如你三次才抛出正面，你付给对手 8 英镑，以此类推。每多抛一次，你的付出增加一倍。

问题是：你的对手应当付给你多少钱才能保证这个游戏是公平的？你在每一轮中的支出将是多少？

在这个问题里我们再次用到平均值。我们可以算出你付给对手钱的平均值，为了保证游戏公平，对手预先付给你的钱应当等于这个平均值。你抛出的结果应当是以下各种情况中的一种：正，正反，正正反，正正正反，……，等等，各种情况对应的概率分别是 1/2，1/4，1/8，……，等等。表 4.2 列出了几种付款额对应的概率。当奖金额倍增时，对应的概率减半。

现在我们来计算你的支出的平均值。根据附录 3 介绍的方法，我们把表 4.2 中的各列中的两个数相乘，然后把所有的乘积加在一起，这个数应该是 1 + 1 + 1 + 1 + 1 + 1 + ……。等一等！你发现哪里不对劲了吗？这个数是无穷大！这意味着，你的对手应当预先支付给你无穷多的钱，因为根据期望值，你将付给对手的钱是无穷多的。

表 4.2：圣彼得堡游戏中付款额与相应的概率

付款额	2	4	8	16	32	……
对应的概率	1/2	1/4	1/8	1/16	1/32	……

这个结论相当可疑，但是我们的计算没有瑕疵，结论确实是这样。可是，即使你像比尔·盖茨一样富有，你也拿不出一笔无穷大的款项呀！其实当钱的数额超过 5 000 亿英镑时，这个数额本身就没有意义了。所以，我们必须为你支出的款项设一个封顶值，当钱款的数额超过封顶值时，就取封顶值。

例如，我们约定 64 英镑为封顶值。就是说，如果你连续抛出 5 个反面，你就不必继续抛了，你的付出额固定为 64 英镑。表 4.2 给出了你在五次以内抛出正面的各种情况。你连续五次都抛出反面的概率是 1/32，此时你的支出是 64 英镑，这两个数的乘积是 2。所以，你的支出的平均值为 1 + 1 + 1 + 1 + 1 + 2 = 7 英镑。结论：你的对手应当预先付 7 英镑。

假设你们在这个约定下进行游戏，玩很多轮。如果你在两次以内抛出正面的频率高于四分之三，你就赚了；反之，你就赔了。就平均值而言，损失和收益相平衡，但是在实战中，盈亏值会有剧烈的变化。

你的对手自己也要计算一下自己应当预先支付多少钱才能保证游戏公平。他的数学知识可能不如你好，所以他需要你的帮助。如果你不像我这样心慈手软，你完全可以这样“指导”他：在每一轮游戏中他的收益可能是 2 英镑、4 英镑、8 英镑、16 英镑、32 英镑及 64 英镑中的一种，所以他的收益的平均值是（2 + 4 + 8 + 16 + 32 + 64）/6 ≈ 21 英镑。这样，如果他预先支付 21 英镑，则游戏是公平的。你可以表现得慷慨一些（也是为了增加游戏对他的吸引力），只要求他支付 15 英镑。在这个游戏中，你有可能输吗？你在利用对手的无知，不过，对待“凯子”何必心慈手软？

我们先把这个陷阱放在一边，再研究一下圣彼得堡游戏。最初的问题是对手需要预先支付的钱是无穷大。如果我们不约

定封顶值，问题同样可以解决。例如，我们调整一下你在各个情况下付给对手的钱数，假如你只用一次就抛出正面，则付出2英镑；假如你用两次抛出正面，则付出4英镑；假如你用3次抛出正面，则付出6英镑。以此类推，每多用一次，则多付2英镑。此时，你的支出的平均值不再是无穷大，而是一个确定的值：4英镑。也就是说，你的对手预先支付4英镑即可保证游戏公平。在这个规则下，你支出的钱数是没有上限的，你**有可能**支出一笔天文数字的巨款——但**仅仅是有可能**，这种事情发生的概率是非常非常低的。事实上，你的支出超过64英镑的概率就已经小得可以忽略不计了。如果你是一个极端谨慎的人，你们可以约定，你的支出以16英镑为封顶值。此时，为了保证游戏公平，你的对手预先支付的钱也应当调整，3.98英镑是最公平的。 53

潘尼游戏

潘尼游戏根据其发明者的名字命名。1969年W·潘尼发明了这种游戏。对于粗心大意的人来说，这种游戏是一个可怕的陷阱。它的特点是看起来对某一个人有利，实际上却对另一个人有利。如果你能找到一个人陪你玩潘尼游戏，你很容易利用对手的无知狠狠地赚一笔钱。

把一枚理想硬币（即正面和反面出现的概率完全相等的硬币）抛30次，可能得到如下结果（为便于阅读，书写时五个一组）：

正反正反反　反正反反反　反反正正反　反正反

反反　正反正反正　正正正反反

潘尼游戏依赖于连续的三个结果的组合。第一个组合是正反正，第二个组合是反正反（第 2、3、4 次抛出的结果），第三个组合是正反反，第四个组合是反反反，等等。一共有八种可能的组合：

正正正　正正反　正反正　反正正　正反反　反正反　反反正　反反反

游戏的规则是你的对手先在八种组合中选一种作为“他的”组合，然后你在剩下的七种组合中选一种作为“你的”组合。双方选定以后，你们中的一个人（或者第三者）连续抛硬币，直到他的组合或你的组合出现时为止。如果他的组合先出现，则他获胜；如果你的组合先出现，则你获胜。

这个游戏似乎对你的对手有利，因为他可以选择任何一种组合，而你只能在他选剩的组合中选择。如果某种组合的出现的频率较高，他有优先选择的权利。无论赌注是多少，你好像都占不到便宜。就像在象棋中先手走总是有利可图，“先下手为强”的原则似乎在很多游戏中有效。可是你知道吗？如果你的对手真的这样想，他可就上当了！

事实是，无论你的对手选择哪一个组合，你都可以针对他的选择做出你的选择，从而使你的选择比他的选择先出现的概率至少多达 2/3！看起来难以置信，是不是？八种组合出现的频率应当是一样的，为什么某一个组合比另一个组合先出现的概率较高呢？其次，即使确实有某一个组合比另一个组合先出现的概率较高，你的对手有优先选择的权利，他还是会先选择

这个有利的组合吗?

我们从简单的例子开始分析。假设对手选择“正正正”，你要小心，一定要选择“反正正”。我可以肯定地告诉你，“反正正”比“正正正”先出现的概率是 7/8。证明如下：

如果前三次就抛出了“正正正”，则你的对手获胜。这种情况发生的概率是 1/8。下面我们集中讨论剩下的 7/8。如果前三次抛出的结果不是“正正正”，你的对手有可能获胜吗?不妨假设他获胜，即“正正正”在某一阶段出现。那么“正正正”的第一次出现是在什么时候呢? 就算是在第 25、26、27 轮出现吧。由于这是“正正正”的第一次出现，所以第 24 轮一定不是“正”，否则第 24、25、26 轮就是“正正正”的第一次出现，与第 25、26、27 轮是“正正正”的第一次出现矛盾。所以第 24 轮一定是“反”，即第 24、25、26 轮一定是“反正正”！这说明“反正正”一定先于“正正正”出现！这与你的对手获胜的假设矛盾。结论是：除非头三次就抛出“正正正”，你的对手必败无疑。他取胜的机会是 1/8，另外 7/8 的机会完全属于你。

对于你的对手选择其它组合的情况，我们不再专门讨论。细致的分析要费一点心思，但是基本原则都是一样的。表 4.3 列出了你需要的全部信息。从表中可见，无论你的对手选择八个组合中的哪一个，他都在你的掌握之中，你至少有 2/3 的机会击败他。

表 4.3：潘尼游戏中的最佳策略与获胜概率

对手的选择	你相应的最佳选择	你获胜的概率
正正正	反正正	7/8

对手的选择	你相应的最佳选择	你获胜的概率
正正反	反正正	3/4
正反正	正正反	2/3
反正正	反反正	2/3
正反反	正正反	2/3
反正反	反反正	2/3
反反正	正反反	3/4
反反反	正反反	7/8

55

惟一需要注意的是千万别让你的对手看到表 4.3。如果你在游戏过程中拿出这张表来选择策略，那无异于自断财源。我来告诉你如何记住这张表的规律：你的选择的最后两项就是对手的选择的前两项，而你的选择的第一项与第三项相反。如此，你就可以保证 2/3 以上的胜率。

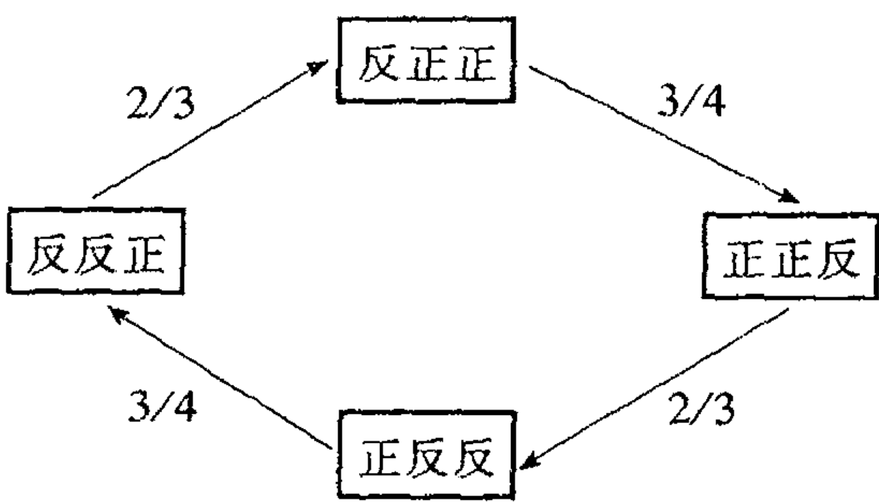
在赔率相等的情况下，你的战果应当相当辉煌。为了增加游戏对他的吸引力，你可以提供优惠的赔率，比如说，你每赢一局他付给你 4 英镑，而他每赢一局你付给他 5 英镑。由于你获胜的次数至少是他的两倍，他还是在你的算计之中。为了确保获胜，应使游戏进行的次数尽量多。如果只玩五六次，说不定你还会输呢！

表 4.3 的第二列构成一个循环："反正正" 优于 "正正反"(3/4)，"正正反" 优于 "正反反" (2/3)，"正反反" 优于 "反反正" (3/4)，"反反正" 优于 "反正正" (2/3)。如图 4.1。

每一个组合优于下一个组合，所以，不存在绝对的最佳组合。但是，相对于你的对手的选择，你总能找到一个最佳组

合。

图 4-1:潘尼游戏中各种获胜组合的关系



潘尼游戏可以有多种变化。我们可以调整一个组合中包含的抛硬币的结果个数，把 3 变成其他数，游戏也是可行的。然而，3 是最佳选择。当我们的选择大于 3 时，在每一轮中为了分出胜负我们通常要抛好多次硬币；当我们的选择是 2 时，又会出现不利的情况：你的获胜策略将失效。当我们选择 2 时，如果你的对手选择组合“正正”，我们有把握打败他，因为我们可以选择“反正”，我们取胜的概率是 3/4；可是如果你的对手选择“反正”，无论你选择什么组合都不会比对手更有利。即使你选择“反反”或“正反”，你也只能保证获胜的概率与对手相等。当我们约定一个组合中包含的抛硬币的结果个数大于 3 时，获胜策略总是存在的，后选择的一方有机可乘。然而，每一轮游戏可能占用很长时间。

在潘尼游戏中，平均每一轮要抛多少次硬币才能分出胜负？表 4.3 没有提供这方面的信息。你认为这种分析是否合理：“一共有八种可能的组合，因此每种组合出现的概率是八分之一。根据附录 3 中的计算原则，如果某个事件发生的概率是 p，则出现一次此事件所需的平均次数是 1/p。所以，每一轮平均需要抛八次硬币。”如果每一种组合在游戏中的出现是 56

相互独立的，以上分析有效。然而事实是，这八种组合是相互影响的，正确的计算方法要复杂得多。如果你多做一些实验，你就会发现，“正正正”和“反反反”出现的频率明显偏低。

对应于表4.3第一列中每一个组合，其所需的抛硬币的次数分别是14、8、10、8、8、10、8、14。第一项（“正正正”）和最后一项（“反反反”）的平均频率比某些组合低将近一倍。另外，前后对称的组合如“正反正”和“反正反”出现的频率也比较低。严格证明这件事要费一点心思。如果你想证明自己的数学才能，我建议你从一个类似的但简单一些的问题入手：如果每个组合的长度不是三而是二，各个组合（“正正”、“正反”、“反正”、“反反”）对应的频率分别是多少？答案是：对于“正反”和“反正”，平均每抛四次硬币出现一次；而对于“正正”和“反反”，平均每抛六次硬币出现一次。下面是对于组合“正正”的分析，没有兴趣了解细节的读者可以略过这一段。如果你能凭自己的才智解决这个问题，我将更加欣慰。

你是否觉得这个答案出乎意料？当我最初见到以上结论时，我感到难以接受。它们与我的直觉矛盾，但是符合逻辑。请注意：在涉及到概率时，我们的直觉通常是错误的。“为了得到‘正正正’，平均需要抛14次硬币”，我们应当如何理解这个频率呢？

你是否经历过这种事情：当你等公共汽车时，半小时里一辆公共汽车也没来，可是随后，接连来了三辆公共汽车！这就是著名的11路公共汽车问题。潘尼游戏中的情况与此类似。我们回到最初的例子：连续抛30次硬币，结果分别是：

正反正反反　反正反反反　反反正正反　反正反反反　正反正反正　正正正反反。

“正正正”第一次出现在25、26、27的位置上。一旦“正正正”出现过一次，它再次出现的概率立刻上升到50%，因为第28次抛硬币出现“正”的概率是1/2，而只要第28次抛硬币出现“正”，在26、27、28的位置上即出现“正正正”。进一步说，由于第29次抛硬币出现“正”的概率是1/2，所以一旦在26、27、28的位置上出现“正正正”，则在27、28、29的位置上出现“正正正”的概率立刻达到50%。这与等公共汽车的例子别无二致——“正正正”倾向于“要来一块来”。

证明：平均每抛六次硬币，将出现一次“正正”的组合。

假设平均每抛X次硬币，将出现一次“正正”的组合。X是一个大于1的数，我们暂且不知道X等于几，但是它是一个确定的值。

在我们抛了若干次硬币之后，可能出现四种情况：情况一：最初的两次结果是“正正”；情况二：最初的两次结果是“正反”；情况三：第一次的结果是“反”。情况一和情况二对应的概率都是1/4，情况三对应的概率是1/2。

对于情况一而言，为了得到“正正”，只需抛两次硬币；

对于情况二而言，为了得到“正正”，需要抛2+X次硬币，因为头两次已经浪费了；

对于情况三而言，为了得到“正正”，需要抛1+X次硬币，因为第一次已经浪费了。

综上所述，$X=2\times(1/4)+(2+X)\times(1/4)+(1+X)\times(1/2)$。

解方程，得

$$X=6$$

假如我们连续抛许多许多次硬币，平均每抛多少次会出现一次“正正正”？答案小于14。需要注意，我们的问题已经变了。我们最初的问题是：第一个“正正正”的出现平均需要抛多少次硬币？它的答案是14。

有四个组合（“正正反”、“反正正”、“正反反”和“反反正”）的第一次出现平均需要抛8次硬币。这四个组合的共同特点是不对称性，即第一个位置的值与最后一个位置不同。这使得它们具备一个性质：第一次出现并不使下一次出现的概率上升。以“正正反”为例，这个组合的最后一位“反”不能成为下一个“正正反”的第一位。所以，每一个“正正反”的出现是相互独立的。另外两个组合（“正反正”和“反正反”）的第一次出现平均需要抛10次硬币。这两个组合的共同特点是对称性，即第一个位置的值与最后一个位置相同。这使得它们具备一个性质：第一次出现使下一次出现的概率上升。以“正反正”为例，这个组合的最后一位“正”可以成为下一个“正反正”的第一位。所以，在一个“正反正”出现以后，只要接着出现一个“反”和一个“正”，就能得到下一个“正反正”。“正反正”和“反正反”也倾向于“要来一块来”。至于“正正正”和“反反反”，情况更明显。以“正正正”为例，这个组合的最后两位可以作为下一个“正正正”的最初两位。所以，在一个“正正正”出现以后，只要接着出现一个“正”，就能得到下一个“正正正”。你是否发现了一个很有趣的规律：一个组合越是倾向于“要来一块来”，它的第一次出现就越困难。

58

正面和反面：哪个出现的次数多？

我们回到最初讨论的问题：连续抛许多次硬币，计算出现正面的总次数和出现反面的总次数。在任何阶段，要么正面出现的次数较多，要么反面出现的次数较多，要么正面和反面出现的次数一样多。第三种情况只有在总次数为偶数时才能出现。平均抛多少次会出现一次正面和反面出现的次数一样多的情况？正面出现的次数较多和反面出现的次数较多这两种情况之间相互转化的频率如何？什么时候会出现最后一次正面和反面出现的次数一样多的情况？如果我们仔细地研究这些问题，我们会发现，正确的答案与常识相去甚远。

以这些问题为基础可以设计出许多赌博游戏。你可以为各种可能的结果设计赔率，然后下注开赌。不过在开始之前，最好对各种概率有清醒的认识，否则你很可能把自己设计了。即使你不关心如何用这方面的知识赚钱，这种问题带来的纯理智上的享受也是很可观的。

你要准备好吃惊。一个了不起的概率论专家——威廉·费勒——就曾被自己的结论震惊。威廉·费勒的概率论著作启发了整整一代人。他的概率论专著的第一版出版于1950年，其目的就是让不熟悉数学的人了解概率。1957年第二版出版，主要的变化就是专门增加一章讨论抛硬币的问题。他在《序言》中说的一段话耐人寻味："关于抛硬币的计算经常表明，那些被大众普遍接受的结论往往是错误的。这些结论与通常的直觉如此之不同，以至于许多训练有素的专家都怀疑其正确性。"再过十年，他又出版了第三版，主要变化仍然在讨论抛

硬币的那一章。第二版的结论没什么错误，但是第三版增加了一些精妙的方法来揭穿某些复杂的陷阱。

数学基础好的读者可以直接阅读费勒的著作。以下我将略过严格的证明，直接给出问题的答案。你也可以用实验来检验给出的结论。我们还是从简单的例子开始：首先分析抛 12 次硬币出现的情况。一方面，12 这个数不太大，这使得计算不会太复杂；另一方面，12 这个数也不太小，这使得概率的原则已经可以生效。

抛多少次硬币出现首次正面和反面出现的次数一样多的情况？

59

如果你不断地抛硬币，将会发生正面和反面出现的次数一样多的情况。当然，你所用的次数可能高于 12。表 4.4 给出了对应于抛 2、4、6、8、10、12 次硬币，首次出现正面和反面出现的次数一样多的情况的概率（这种情况的首次出现称为“最初平衡点”）。为了使正面和反面出现的次数一样多，你所需要的次数可能高于 12，这个事件的概率是 231/1 024，约等于 23%。请注意：表 4.1 与表 4.4 不同。表 4.1 给出的是“正面和反面出现的次数一样多”的概率，表 4.4 给出的是“首次出现正面和反面出现的次数一样多”的概率。

表 4.4：出现首次正面和反面出现的次数一样多的情况的概率

抛硬币的次数	2	4	6	8	10	12
对应的概率	1/2	1/8	1/16	5/128	7/256	21/1024

如果你要求正面和反面出现的次数一样多的情况是首次出现，你需要尝试许多次才能遇到一次；相反，如果你只要求正面和反面出现的次数一样多，而不在乎以前是否已经出现过，则你需要的次数就少得多。关键在于，如果你抛N次硬币以后遇到了正面和反面出现的次数一样多的情况，那么你再抛两次硬币，正面和反面出现的次数依然一样多的概率明显上升——概率是二分之一。有点像我们刚刚讨论过的等公共汽车的情况，我们期待的事情倾向于“要来一块来”。不同之处在于，随着N的增加，正面和反面出现的次数一样多的概率递减，而且趋近于无穷小——无论我们是否要求“首次出现”。这意味着，当N足够大时，我们期待的事情可能永不发生。

频率倒转

首先我们定义何谓“频率倒转”。

连续抛硬币，如果某一阶段符合以下条件：

(1) 在某一次抛硬币（第N次）之后，正面和反面出现的次数一样多；

(2) 上一次抛硬币（第N-1次）之后，正面或反面中的某一个出现的次数较多，而下一次抛硬币（第N+1次）之后，另一个出现的次数较多。

这种情况称为频率倒转。

由于正面和反面出现的次数一样多的情况必须出现在抛偶数次硬币之后，所以频率倒转的情况必须发生在抛奇数次硬币之后。为避免混乱，我们约定：在第N+1次抛硬币之后才认为频率倒转已经发生。

在继续阅读之前，请你先考虑两个问题，看看你的答案是否与我相同。

问题一：连续抛 101 次硬币，你认为会出现多少次频率倒转？你可以在三个候选答案中选择一个：

a. 不超过四次；

60

b. 五至九次；

c. 不低于十次。

问题二：连续抛 20 001 次硬币，你认为会出现多少次频率倒转？

问题二的答案是 56。

连续抛 N 次硬币，用 x 表示发生的频率倒转的次数。x 可以是 0，1，2，3，……。对应于每一个 x 的可能值，相应的概率是多少？我可以肯定地告诉你：x = 0 的概率高于 x = 1 的概率；x = 1 的概率高于 x = 2 的概率；x = 2 的概率高于 x = 3 的概率……。随着 x 的取值的增加，相应的概率递减。你可以观察篮球赛中的情况：在一场比赛中会出现多少次比分交替领先的情况？出现 0 次的概率最大，出现 1 次的概率次之，出现 2 次的概率再次之，等等。随着次数的增加，概率递减。

以上结论都是根据一个著名的公式计算而来的。非常遗憾，这个公式的推导过于专业化，我不能在这本书中介绍。如果你不满足于简单地接受我的结论，你可以参考费勒的专著。如果你对自已的数学才能有足够的自信，你也可以尝试独立地证明它。这个公式说明：频率倒转的次数与正面出现的总次数有关。这个结论看起来非常奇怪，因为频率倒转依赖于具体的过程，而正面出现的总次数依赖于对整体的统计。公式的严格

表述如下：

> 频率倒转的次数为 x 的概率是正面出现的次数比反面出现的次数多 2x+1 的概率的 2 倍。

对于抛 101 次硬币的例子而言，x=0（即频率倒转没有发生）的概率等于正面出现的次数比反面出现的次数多 1 次的概率的 2 倍。（也就是说，出现 51 次正面和 50 次反面。）x=1（即频率倒转发生 1 次）的概率等于正面出现的次数比反面出现的次数多 3 次的概率的 2 倍。（也就是说，出现 52 次正面和 49 次反面。）类似地，出现 2 次频率倒转的概率等于出现 53 次正面和 48 次反面的概率的 2 倍，以此类推。显然，频率倒转没有发生的概率最大。表 4.5 给出了最初的几个概率值。注意：以上分析不仅对抛 101 次硬币的例子有效，其方法是通用的。 61

表 4.5：抛 101 次硬币，频率倒转的次数对应的概率

频率倒转的次数	0	1	2	3	4	5
对应的概率	0.158	0.152	0.140	0.125	0.107	0.088
频率倒转的次数	6	7	8	9	10	10 以上
对应的概率	0.069	0.052	0.038	0.027	0.018	0.026

下面我们看看问题一的答案。频率倒转不超过四次的概率是 68%，频率倒转五至九次的概率是 27%，频率倒转不低于十次的概率仅仅在 4% 到 5% 之间。事实上，频率倒转不超过 1 次的概率有 31% 之高！如果你没有猜对答案，不必伤心，我

也没猜对。大多数人都把频率倒转的次数远远地高估了。如果你能找到一个人，和他赌频率倒转的次数不超过四次，平均每三轮中你可以赢两轮。（当然，条件是你的对手没读过这本书！）

为了利用频率倒转的知识赢钱，你应该学会选择多大的频率倒转次数下注。抛 N 次硬币，可以确定一个值 x，你赌频率倒转的次数不超过 x，而对手赌频率倒转的次数超过 x，则双方的胜率大致相同，这个 x 的值即为“收支平衡点”。如果你选择一个比收支平衡点大的值 A，赌频率倒转的次数不超过 A，游戏就是对你有利的。收支平衡点可以这样估计：对于抛 N 次硬币的情况，首先算出 N 的平方根，而后除以 3。如此得到的结果和精确值很接近。为了确保万无一失，你可以把这个数增大一点——增大 20%就很理想。最后你得到一个确定的数 A，然后你可以打赌频率倒转的次数不超过 A。例如，当 N = 101 时，N 的平方根约等于 10，除以 3 的结果略大于 3，再增大 20%得到 4。于是，你可以赌频率倒转的次数不超过 4。如果 N = 251，则你可以赌频率倒转的次数不超过 5。你是否相信，不发生频率倒转的概率高达 10%！

最后平衡点

抛若干次硬币，也许正面出现的次数多，也许反面出现的次数多。如果你不停地抛下去，不可避免地会出现正面和反面出现的次数一样多的情况。（严格地说，正面和反面出现的次数一样多的情况始终不出现的可能性是存在的，然而，当抛硬币的次数较多时，这种情况的概率非常接近于 0。）抛 N 次硬

币（N为偶数），最后一次出现正面和反面出现的次数一样多的时刻称为“最后平衡点”。用X表示最后平衡点所对应的抛硬币的次数。当N=200时，你认为最后平衡点对应的X值应当是多少？也许在整个过程中正面出现的次数和反面出现的次数始终不相等，此时我们认为X=0，因为在抛第一次硬币之前正面出现的次数和反面出现的次数是一样多的——都是0。也许在200次结束以后正面出现的次数和反面出现的次数恰好相等（都出现100次），则最后一轮就是最后平衡点，此时X=200。一般地，X可以取0到200之间的任何偶数。X可以是101个数中的一个，你认为X出现在什么范围内的概率最大？接近于0？接近于100？还是接近于200。在继续阅读之前，请做出你的猜测。 62

惟一使X=0的情况即在全部200次抛掷过程中一次正面和反面出现的次数相等的情况也没出现，看起来这个概率是很小的。但是如果X出现在0到100之间，没什么好奇怪的。你在看篮球比赛时可能已经注意到，有二分之一的可能后半场不出现比分相等的情况。X出现在0到100之间的概率与X出现在100到200之间的概率相等。

这说明，如果抛完100次硬币以后，正面出现的频率较高，那么有二分之一的可能在此后的全部过程中正面出现的次数始终领先。假设有两个人被恐怖分子囚禁，他们没有书，没有象棋，没有扑克，只能靠一枚硬币打发时间。他们每天抛5小时，每小时抛600次，坚持抛1年。如果到6月底时正面出现的次数高于反面出现的次数，则有二分之一的可能在以后的半年中正面出现的次数始终高于反面出现的次数。

这个理论适用于体育比赛。在一场篮球赛中，我们可以假设在每一时刻对阵双方得分的机会是均等的，此时双方的得分

相当于抛硬币时统计的正面和反面出现的次数。假如在上半场结束时某个队领先，有二分之一的可能这个队把优势保持到终场。在足球联赛中，如果有一个队在赛事进行到一半时积分领先，这个队有多少机会把优势保持到底？这个问题要复杂得多，因为参赛的球队不只两支。

如果你理解了以上原理，计算 X 取各个值的概率变得很简单。X 取值的概率分布具备严格的对称性，即 X = 190 的概率等于 X = 10 的概率，X = 198 的概率等于 X = 2 的概率。一般说来，对于整数 n（n < 200），X = n 的概率等于 X = N - n 的概率。（这个令人惊异的规律是由三个数学家——戴维·布莱克威尔和他的两个加利福尼亚同事——在 1964 年无意中发现的。当时他们想检验一台计算机的运行结果是否可靠，方法是让计算机计算两组数，这两组数中位置相同的一对数的和应当等于 1。计算机确实算出了两组数，而且两组数中位置相同的一对数的和果然等于 1，但是这两组数与他们的预期不同。他们开始研究计算机给出的两组数的含义，结果导致了本原理的发现。他们给出了一个极其简洁的证明。）

63

只要我们知道了 X 在 0 到 100 之间取值的概率分布，就知道了 X 在 100 到 200 之间取值的概率分布。在 0 到 100 之间，随着 X 取值的增加，对应的概率值递减。这说明，X 等于 0 或 200 的概率最大，而 X 等于 100 的概率最小！

假如你利用最后平衡点和别人打赌，你应当赌最后平衡点出现在最初或最后。此时你获胜的概率最大，在 5% 到 6% 之间。这个概率不够大，但是如果你赌其他结果，概率更低。毕竟你是在 101 种可能的结果中赌 1 种，5% 到 6% 的概率已经很大了。假如别人已经选择了最初和最后，那么你的最佳选择就是赌第 2 轮和第 198 轮。X 的取值在 {0，2，198，200} 这四

个数之中的概率高达17%。相反，如果你赌100附近的数，你的前途非常暗淡。X的取值在{90，92，94，……，108，110}这11个数之中的概率只有7%。

这个结论让你意外吗？其实，就连威廉·费勒这样的专家也对这个结果吃了一惊。多数人认为，最后平衡点出现在中间的概率最大，当然你已经知道，这是大错特错。如果你想把关于最后平衡点的知识应用于赌博实战，可以把抛硬币的总次数削减到20，这样每一局不会占用太长时间。表4.6给出了最后平衡点的各个位置对应的概率，你可以利用这个表下注。

概率值从两端向中间递减。一般地，如果抛2n次硬币，则X=0和X=2n对应的概率最大，这个值接近于$1/\sqrt{(n\pi)}$。

我们再次证明，关于概率的直觉通常是错误的。

表4.6：抛20次硬币，最后平衡点的位置对应的概率

最后平衡点的位置	2(20)	2(18)	4(16)	6(14)	8(12)	10
对应的概率	0.176 2	0.092 7	0.073 6	0.065 5	0.061 7	0.060 6

64

排队问题

在邮局、超市和火车站，我们经常需要排队。你是否多次遇到这种情况：两条队看起来一样长，不知道该排哪条队好？你是否和我一样随便选择一条队？你是否和我一样最后总是发现自己排错了队！

排队问题可以用抛硬币的模型分析。假设你加入一条队，另一个人——你的对手——与你同时到达，而他选择了另一条

队。有一个精灵在抛硬币，如果抛出正面，则你前进一步；如果抛出反面，则你的对手前进一步。经过一段时间以后，谁会领先呢？根据我们以前的分析，你们中的某一个在一长段时间内保持领先是正常的，而频繁地交替领先概率极低。至于你们中的哪一个比较幸运，完全是随机的。为什么我们总是觉得自己幸运的时候比不幸的时候少呢？那是因为我们的记忆是有选择的。我们更容易记住自己的队移动得较慢的情况。

假如一开始你的对手就领先于你，你最后超过他的机会有多大？很遗憾，概率很低。不过别伤心，你的对手面临同样的问题！

小　结

所有讨论基于一个假设：硬币是公平的，即正面和反面出现的概率完全相等。不过即使硬币本身不是完全公平的，我们也很容易“设计”一个理想硬币。

圣彼得堡游戏的胜负取决于你用多少次抛出正面。为了使游戏公平，需要调整双方的赔率，规定好你每多抛一次付多少钱，以及对手预先支付多少钱。

潘尼游戏的胜负取决于哪个组合先出现。这种游戏对后选择的一方有利。无论对手选择了哪个组合，你都可以相应地选择一个组合，确保你获胜的概率不低于三分之二。选择出现的频率较高的组合比较有利。一个组合越是倾向于“要来一块来”，则其出现的频率越低。

连续抛许多次硬币，计算正面出现的总次数和反面出现的总次数。虽然从理论上说这两个数应当相等，可是这两个数完

全相等的概率是很低的，至少比大多数人的估计低得多。频率倒转发生的次数为0的概率高于任何确定的值。通常在一段很长的时间不出现正面和反面出现的次数相同的情况，而在另外一段时间，接连出现这种情况。最后平衡点出现在最初和最后的概率最大，而出现在中间的概率最小。如果不经过仔细的计算，直觉经常把我们引向错误的方向。

练习题

1. 利用斯达灵公式计算：抛50次硬币得到的正面和反面次数相等的概率是多少？抛250次概率是多少？抛1 000次概率是多少？

2. 修改潘尼游戏的规则，每个人选择一个由四个结果组成的组合。假如你的对手选择了"正正正正"，你的最佳选择是什么？你获胜的概率是多少？假如你的对手选择了"反反正正"，你的最佳选择是什么？

3. 比较表4.1和表4.4，计算两个表中对应的概率值之间的比率。你发现其中的规律了吗？

4. 假设你已成功地解决了问题1和问题3，则最初平衡点出现在第50次的概率是多少？第250次呢？第1 000次呢？ 66

5. 抛1 001次硬币，出现多少次频率倒转的概率最大？最后平衡点出现在哪里的概率最大？

6. 抛1 001次硬币，出现的频率倒转的次数

a) 不低于30；

b) 在20至29之间；

c) 在10至19之间；

d) 不超过 9。

你认为这四个判断中哪个概率最大？哪个概率最小？

67 抛 10 001 次硬币的情况如何？

第五章 骰　　子

人类玩骰子的历史源远流长。5 000 年以前古埃及人就开始玩骰子。在特洛伊战争期间，古希腊人发明了多种骰子游戏来排遣枯燥的军旅生活。古罗马帝王也喜欢骰子：奥古斯都曾在书信中纪录自己玩骰子的经历，克劳迪亚斯甚至有一块专门用来掷骰子的棋盘，放在马车里供旅途中使用。据 F.N. 戴维在《游戏，上帝和赌博》一书中介绍，四足动物的踝骨形状很规则，适合制造骰子。古代的骰子通常用这种骨头制造。

2 500 年以前人们就已经知道，只存在五种外形完全对称的多面体，即所谓的“正多面体”。所以，标准的骰子只能是五种形状中的一种。除了最常见的正方形骰子以外，另外四种是：

金字塔型骰子，四个面，每个面都是三角形；

钻石型骰子，八个面，每个面都是三角形；

正十二面体骰子，十二个面，每个面都是五边型；

正二十面体骰子，二十个面，每个面都是三角

形。

这四种骰子都可以买到，但是在赌博中使用的骰子几乎都是正方形的。考古学家曾发现 14 个面、18 个面和 19 个面的骰子，不过这些骰子的形状不是对称的。在本章中，如无特别声明，则我们讨论的骰子都指公平的正方形骰子：六个面，分别标以 1 至 6 的号码，每个号码出现的机会相等，每次投掷的结果相互独立。

一枚骰子的情况

68 在统计学上，我们如何判定一枚骰子是否公平？与判断大英帝国彩票的获奖号码是否随机一样，问题取决于两个因素：你可以容忍多大的与绝对公平的结果的偏差，以及你获得数据的丰富程度。也许你这样考虑：掷 120 次骰子，如果每个号码出现的次数都是 20，则认为骰子是公平的。如果你以此为标准，几乎所有的骰子都会被你淘汰。我们必须考虑这个问题：距离平均值的多大的偏差是可以接受的？先考虑这个问题：如果有一枚做过手脚的骰子，每一次掷骰子时有三个号码出现的概率是 0.147，而另外三个号码出现的概率是 0.187，我们是否可以通过检验发现它的毛病？

我们可以用计算机模拟检验过程。我设计了一个计算机程序，这个程序随机地产生从 1 到 6 的号码，其中 1、2、3 产生的概率是 0.147，而 4、5、6 产生的概率是 0.187。每次产生的号码是相互独立的。每次运行产生 120 个号码，共运行 4 次。统计每个号码出现的频率，记录在表 5.1 中。

如果我们只掌握了120次骰子的数据，我们不可能发现检验结果中的问题，即使我们事先知道骰子是有毛病的。比如说，我们只有第一轮检验的数据，我们会认为号码5出现的频率较低，但实情恰好相反。第二轮的结果显示，这个骰子是没有问题的。在第三轮中，可以发现号码2和号码3出现的频率明显偏低，但是号码1看起来是正常的。第四轮的结果显示了号码1有问题，但是另外五个号码看起来频率正常。

这说明，为了判断一枚骰子是否公平，需要大量的数据。仅仅120次实验是远远不够的。事实上，即使我们把四轮检验的数据合并在一起，得到480次实验的结果，我们的数据仍然不够丰富。显然，号码五出现的频率低于1/6，但事实上5是出现的概率较高的号码。附录4讨论了更多的细节。

表5.1：计算机模拟不公平的骰子，四轮检验的结果

	号码1	号码2	号码3	号码4	号码5	号码6	总计
第一轮	19	21	18	20	13	29	120
第二轮	15	20	16	26	22	21	120
第三轮	25	14	11	22	21	27	120
第四轮	12	21	24	21	23	19	120

69

我们需要多少次实验才能确保数据足够丰富呢？100年前一个统计学家——W.F.R.威尔顿研究过这个问题。在《大英百科全书》第11版关于“概率”的一节中纪录了威尔顿的研

究结论。威尔顿用 12 枚骰子做了 26 306 次实验，发现号码 5 和 6 出现的概率高于平均值。当然，我们没有确切的根据怀疑你的骰子一定有问题，但是你必须意识到，想发现一枚骰子的微小但至关重要的偏差需要做很多工作。

即使在卢度棋这样简单的游戏中，概率知识也是有用的。[①] 当你掷出一个“6”时，你必须决定是让棋盘上的活动棋子移动还是加入一个新的活动棋子。如果你选择前者，你应该知道需要再掷多少次骰子才能得到下一个“6”。

平均而言，你需要再掷六次。你不得不掷 12 次以上才能得到“6”的概率是多少？这意味着连续 11 次的结果都不是“6”。每一次结果不是“6”的概率是 5/6，所以掷 12 次以上才能得到“6”的概率是 $(5/6)^{11}$，约等于 13.4%。类似地，你不得不掷 20 次以上才能得到“6”的概率是 3%。所以，当你的活动棋子距离终点很近时，你等待下一个“6”是要冒一定风险的，你有可能在等待过程中错过很多机会。因为每次掷骰子的结果是相互独立的，所以你有可能不得不等很长很长时间。然而，平均而言，你需要等六次。

卢度棋中的许多场合需要用到概率知识。不过卢度棋只是一个简单的游戏。有些骰子游戏比卢度棋复杂得多。我们首先介绍西洋双陆棋。

---

① 卢度棋是一种用骰子玩的棋盘游戏，玩家通过掷骰子决定自己的棋子走几步，类似于中国人玩的“大富翁”或“强手棋”。通常四个人分两伙对阵，每个人有四个棋子。一方的所有棋子达到终点则获胜。通常由小孩玩，也称“印度棋”。——译者注。

两个骰子的游戏——西洋双陆棋

本书不准备详细介绍西洋双陆棋的基本规则。[①] 双陆棋的历史可以追溯到5 000年以前，但是现代双陆棋规则直到1925年才确定下来。在任何时刻，一个玩家可以要求“加倍”（即 70 赌注加倍），而另一个玩家可以决定是否接受加倍。如果接受，则赌注加倍，游戏继续；如果不接受，则自动认输，游戏结束。如果一个玩家的加倍要求被另一个玩家接受，则后者可以随时要求“再加倍”，由前者决定是否接受再加倍。当然，接受再加倍的一方还有权要求“再再加倍”，依此类推。

当对方要求加倍时，在什么情况下你应当接受加倍？你需要计算各种情况发生的概率，以及你接受或不接受加倍的得失。如果你接受加倍，你额外获得要求再加倍的权利，不过这点优势暂时可以忽略。

假设当下的赌注为1，而如果游戏继续进行你获胜的概率是x。如果你不接受加倍，你输掉1。如果你接受加倍，赌注变成2。你获胜和失败的概率分别是x和1－x，所以平均你将赢得

$$2\times x+(-2)\times(1-x)=4x-2$$

---

① 双陆棋由两人对阵，分为黑白双方，每方有15枚棋子，分布在棋盘上。双方轮流掷骰子移动棋子。一次掷两枚骰子，每枚骰子的号码各决定自己的一枚棋子的移动，也可以让自己的同一枚棋子移动两次。任何一方的全部棋子达到终点则获胜。双陆棋的规则很复杂，游戏既要求技巧也凭借运气。20世纪后期双陆棋在世界范围内流行，在东地中海国家尤其兴盛。——译者注。

比较这个值和 -1，我们发现当 $x=1/4$ 时二者是相当的。所以结论是：当你获胜的概率低于 1/4 时，你不应接受加倍；当你获胜的概率高于 1/4 时，你应当接受加倍。不论赌注是否已经被加倍过，此结论都有效。

刚才我们没有考虑接受加倍的额外优势：获得要求再加倍的权利。如果在游戏的早期，你接连掷出有利的号码，或者局势向着你喜欢的模样发展，你可以通过要求再加倍向对手施加压力。所以要求再加倍的权利肯定是有价值的，但是其价值很难量化计算。所以，当你获胜的机会比 1/4 略低一点时你还是应当接受加倍，不过你要冒着输 2 倍的风险。

更困难的问题在于什么情况下你应当要求加倍。如果你获胜的机会低于 1/2，要求加倍显然不是好主意，只有在局势对你有利时加倍才是有利可图的。加倍可以使你潜在的盈利倍增，但同时，你已经把要求下一次加倍的权利让给了对手。如果此后游戏又向对他有利的方向发展，他可以用再加倍向你施加压力。所以在加倍以前一定要考虑周详，在游戏的早期，只有获胜的机会很大时才应要求加倍，而在游戏的后期，只要有微弱的优势就可以要求加倍。如果你的优势很明显，你的对手很可能爽快地认输，以避免更大的损失。不过双陆棋本来就是冒险的游戏，什么情况都有可能发生。如果你要求加倍的经历很多，对手在明显的劣势下接受加倍的情况时有发生，反败为胜也不奇怪。

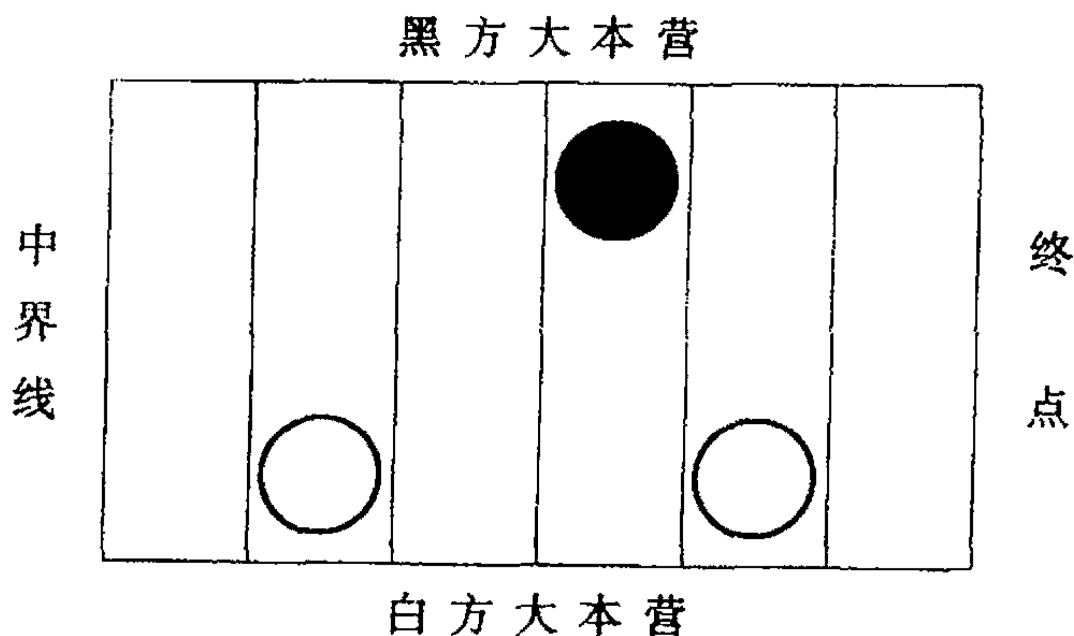
下面的例子可以用来检验你对概率的理解。以下例子来自爱德华·派克欧的《游戏与赌博中的数学》一书。

例一：

游戏接近终局。现在轮到你掷骰子，而且你有权要求加

倍。如图 5.1 所示：你还剩两枚棋子未到达终点，分别距终点 2 格和 5 格。对方只剩一枚棋子未到达终点，距终点仅 3 格。如果你在这一轮不能取胜，对手必然获胜。[①] 此时你是否要求加倍？

图 5.1



我们可以通过查数解决这个问题。两枚骰子可能掷出 36 种结果，其中的 17 种将导致你的失败。(只要你掷出的结果中包含一个 1，你必输无疑。这种情况对应 11 种可能；另外的 6 种结果也导致你的失败：3－2，2－3，4－2，2－4，4－3，3－4。) 所以，你获胜的概率是 19/36，对手获胜的概率是 17/36。局势对你有利，所以你可以要求加倍。从对手的角度看，他获胜的概率高于 1/4，所以他会接受你的加倍。如果你不要求加倍，你本局的平均收益为

① 两枚骰子至少掷出 2。根据双陆棋的规则，如果两枚骰子得到的号码一样，则获得二倍的移动机会。即使对手掷出的是两个 1，他也可以移动两个格，保证达到终点。——译者注。

(1) × (19/36) + (-1) × (17/36) =1/18。

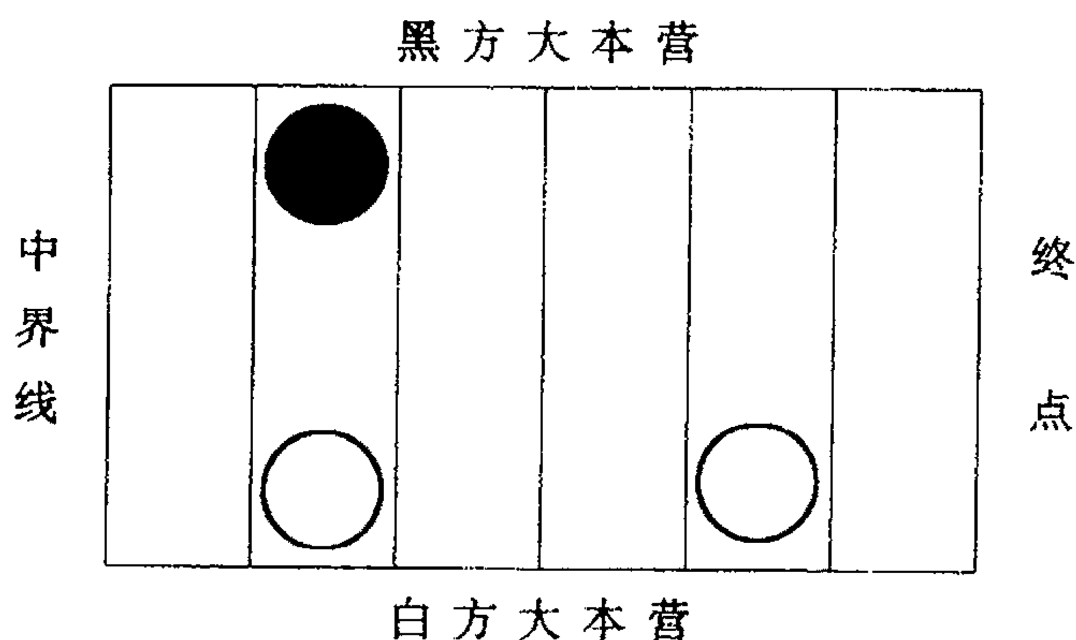
如果你要求加倍，你本局的平均收益增加一倍。所以你应当要求加倍。

例二:

如图 5.2 所示，局势与例一相同，惟一的差别是对手的棋子距离终点更远——还差 5 格。此时你应要求加倍吗?

你是不是觉得这个问题有点蠢? 现在的形势比例一更有利，在例一中我们都要求加倍，此时岂非更该要求加倍? 可是正确答案恰好相反：此时不该要求加倍!

图 5.2



我们来仔细分析这个局面。如果游戏继续进行，有三种可能结果:

1. 你掷一次骰子，立即获胜;

2. 你掷一次骰子，没有获胜。他开始掷骰子，并且获胜;

3. 你掷一次骰子，没有获胜。他开始掷骰子，也没有获胜。

假如你不要求加倍，赌注为 1。在例一中我们已计算过，情况 1 发生的概率是 19/36。如果你掷一次骰子而没有获胜（这个概率是 17/36），则轮到对手掷。他可能掷出 36 种可能，其中 5 种不能使他立刻到达终点（1－1，2－1，1－2，3－1，1－3），另外 31 种使他取胜。在你不能立即取胜的前提下，他取胜的概率是 31/36。因此，情况 2 发生的概率是（17/36）（31/36）＝527/1 296。如果情况 1 和情况 2 都不发生，则情况 3 必然发生，此概率为 85/1 296。在情况 3 下，又轮到你掷骰子，你获胜的概率极高。此时你可以要求加倍，而对手一定不敢接受加倍，所以在情况 3 下，你将取胜。综上所述，在不要求加倍的情况下你的平均收益是

（1）×（19/36）＋（－1）×（527/1 296）＋（1）×（85/1 296）＝121/648。

假如你要求加倍，而对手接受加倍，你的平均收益是多少？你有 19/36 的概率立即获胜。如果你不能立即获胜（概率是 17/36），则轮到他掷骰子，而且他有权要求加倍。由于他取胜的机会不低于 31/36，他一定会要求加倍，而你将认输。所以如果你要求加倍，你的平均收益是

（2）×（19/36）＋（－2）×（17/36）＝1/9。

1/9 小于 121/648，所以你不应要求加倍。

例一和例二有什么根本性的差别？尽管在例二中获胜的概率更大，为什么在例一中应当要求加倍而在例二中反而不应当要求加倍？如果你坚持一个教条——“只要局势对我有利，就

要求加倍”，那是危险的。即使接近终局时也不应如此轻率。关键在于，一旦你要求加倍，下一次要求加倍的权利就让渡到对方手中。

如果你能在游戏中随时计算出自己获胜的概率，那无疑是对你有好处的。不过你完全没有必要计算得非常精确。最实用的办法是经常判断自己获胜的概率是否高于四分之一。

就本质而言，双陆棋类似于赛跑，所以双方的棋子距离终点有多远是一个重要的参数。比赛开始时你的 15 枚棋子距终点的距离总和为 167 格，平均每抛一次骰子你的棋子可以前进 8 格多一点（当两枚骰子的号码相同时你可以走 2 倍的距离），而每一枚棋子的移动范围在 0 到 24 之间。在游戏中有很多不确定的因素，某一次掷骰子的结果的一点微弱变化都可能对胜负产生巨大的影响。在某一轮你可能遥遥领先，可是下一轮你就有可能居于劣势。

很多人认为：“既然前景如此难以预料，我们这样耗神地计算最佳策略有什么用呢？”有时候确实没什么用，然而严肃的参与者的信条是“不放过任何增加优势的机会”，不论这个机会多么微小。从长远看这种态度是绝对正确的。不妨回顾一下 1974 年的世界杯足球赛：某一个阶段的赛程取决于上一个阶段各个队与扎伊尔队的比赛结果。南斯拉夫队与扎伊尔比赛时，在已经获得巨大优势的情况下依然狂攻不止，最终以 9 比 0 大胜。苏格兰队与扎伊尔比赛时，则满足于 2 比 0 的战绩。苏格兰队是失败者。

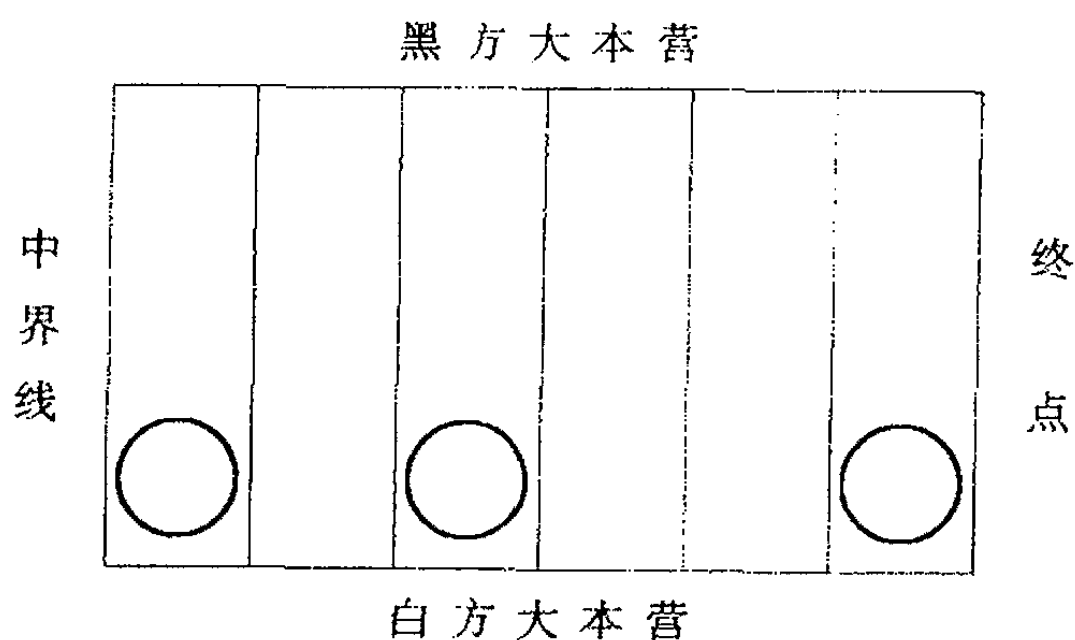
例三：

如图 5.3 所示，你还剩 3 枚棋子未到终点，距终点的距离分别是 1 格、4 格和 6 格。此时你掷出两个 1，即你可以让你

的棋子移动四个1格。有一个棋子（距离终点1格的棋子）可以到达终点，另外两枚棋子应如何安排？你有三个选择：1. 距终点6格的棋子前进2格；2. 距终点4格的棋子前进2格；每个棋子向前移动1格。哪个选择可以使你下一轮获胜的机会更大？

下一轮你掷骰子的结果有36种可能。为了比较以上三个选择的优劣，你只需算出在每一种选择下有多少种掷骰子的结果可以使你立刻获胜？对应于三个选择，答案分别是15、19和17。所以比较而言，选择2最佳。

图5.3

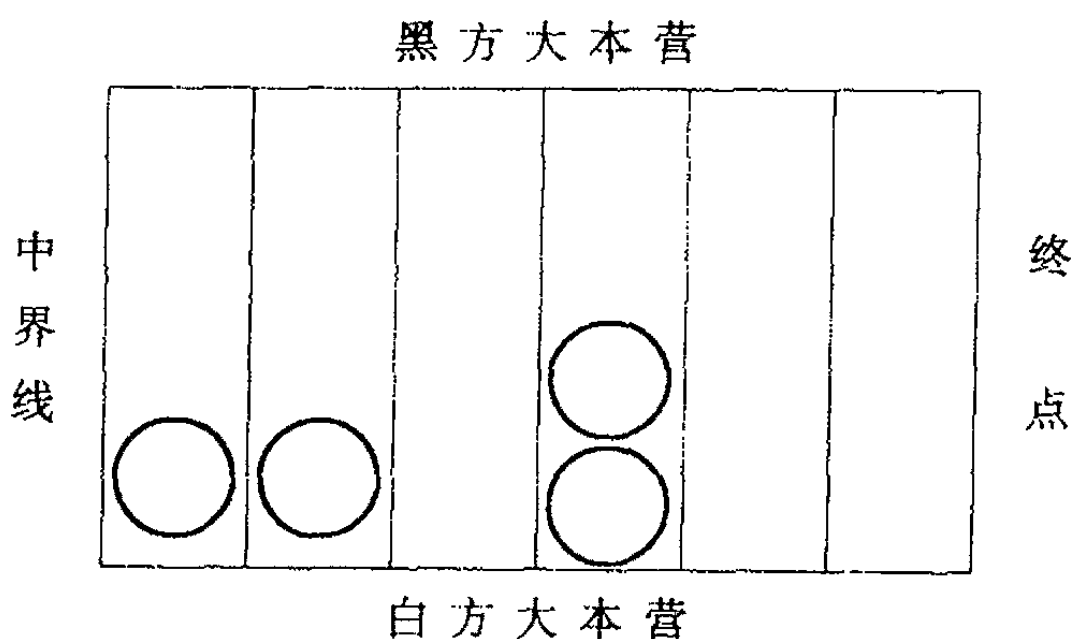


例三是一个非常典型的局势，在实战中你会经常遇到类似的局面。一个高超的玩家能记住许多局势，用以指导自己的实战。如果你经常研究终局形式，你会发现许多重要的规律，这些规律可以帮助你临场时发现最佳策略。例三表现了两个重要的规律：1. 尽量使较多的棋子首先到达终点；2. 尽量使自己的棋子在棋盘上均匀地分布。在多数情况下这两条规律是非常好的指导原则，但它们也有失效的时候，比如例四。

例四：

如图 5.4 所示，你还剩四枚棋子没到达终点，其中两枚距终点 3 格，另外两枚分别距终点 5 格和 6 格。你掷出的骰子的号码是 2 和 1。如何行棋使你在下一轮获胜的机会最大？

图 5.4



如果你遵循“尽量使较多的棋子首先到达终点”的原则，你应当首先把一枚距终点 3 格的棋子送到终点。此时你在棋盘上还剩 3 枚棋子，分别距终点 3 格、5 格、6 格。下一轮，你只有掷出 6－6 或 5－5 才能获胜。如果你换一个策略，让距终点 5 格的棋子前进 1 格，让距终点 6 格的棋子前进 2 格，此时你在棋盘上还剩四枚棋子，其中两枚距终点 3 格，另两枚距终点 4 格。请注意，这个策略同时违反了“尽量使较多的棋子首先到达终点”和“尽量使自己的棋子在棋盘上均匀地分布”这两条原则，但是你获胜的机会却增加了。你在掷出 6－6、5－5 和 4－4 的情况下都能获胜。遵循原则的策略的胜率是 2/36，不遵循原则的策略的胜率是 3/36。后者的概率比前者高 50%。

某些类似于例四的场合会使我们的原则失效。请注意例四的一个特征：你可以选择留下4枚棋子，也可以选择留下3枚棋子，无论如何选择，只有掷出两个相同的号码才能立即获胜。如果你可以在留下3枚棋子和留下2枚棋子之间选择，毫无例外选择后者是正确的。你留下的棋子个数是奇数还是偶数，这个差别至关重要。 75

假设白方还剩六枚棋子而黑方还剩五枚棋子，双方棋子距终点的总距离相等。如果此时轮到白方行棋，则局势对白方有利，因为平均而言双方都需要掷三次完成。假设换一种情形，白方还剩五枚棋子而黑方还剩四枚棋子，双方棋子距终点的总距离相等。如果此时轮到白方行棋，则局势对黑方有利，因为平均而言白方需要掷三次完成，而黑方需要掷两次完成。

例五：

考虑以下五种残局局势：

(1) 还剩2枚棋子，距终点距离分别是6格和4格；

(2) 还剩3枚棋子，距终点距离分别是5格、4格和1格；

(3) 还剩3枚棋子，距终点距离分别是6格、2格和2格；

(4) 还剩4枚棋子，距终点距离分别是4格、3格、2格和1格；

(5) 还剩5枚棋子，距终点距离分别是3格、2格、2格、2格和1格；

这五种局势的共同点是全部棋子距终点的总距离相等，都是10格，而主要差别在于棋子数目不同。在每种情况中，掷一次（或两次以内）骰子结束游戏的概率是多少？在继续阅读之前，请独立考虑次问题，并排出优劣次序。

显然，局势（1）的机会最好，而局势（5）的机会最差。局势（2）、（3）、（4）的优劣则不大明显。想知道确切的结论需要仔细的运算。我认为最好的计算方法是分两步考虑：第一步，计算掷一次骰子即可获胜的概率。首先考虑掷出的两枚骰子号码相同的情况（共6种），然后考虑另外的15种情况（需要注意每种情况对应两种可能结果，因为4-3和3-4是等效的）；如果掷一次没有获胜，继续考虑第二步：在第一次掷骰子得到的号码相同的条件下比较局势（2）、（3）、（4）的优劣。

结论细节如下：对应于掷一次骰子的36种可能结果，有8种结果使局势（1）获胜，有3种结果使局势（2）获胜，有4种结果使局势（3）获胜，有3种结果使局势（4）获胜，有0种结果使局势（5）获胜。显然，剩两枚棋子的局势优于剩三枚棋子，剩四枚棋子的局势优于剩五枚棋子。对应于掷两次骰子的$36^2=1\ 296$种可能结果，有1 203种结果使局势（1）获胜，有1 187种结果使局势（2）获胜，有1 047种结果使局势（3）获胜，有1 127种结果使局势（4）获胜，有335种结果使局势（5）获胜。所以，如果目标是两次以内获胜，五种局势的优劣顺序是（1）、（2）、（4）、（3）、（5）。

76

用百分比表示的结论是：对于局势（1），一次获胜的概率是22%，两次以内获胜的概率是93%；对于局势（2），一次获胜的概率是8%，两次以内获胜的概率是92%；对于局势（4），一次获胜的概率同样是8%，但两次以内获胜的概率只有87%；局势（4）优于局势（3），看起来有点奇怪。对于局势（3），一次获胜的概率是11%，两次以内获胜的概率是82%；对于局势（5），一次获胜的概率是0，两次以内获胜的概率是26%。

在双陆棋中另外一个经常需要考虑的情况是“孤子”。[1]如果你必须使自己的某些棋子成为孤子，你应当知道你的孤子被对方俘虏的概率是多少。

表 5.2：各种距离对应的俘虏对方棋子的概率

距对方棋子的距离	俘虏对方棋子的概率
1	11/36
2	12/36
3	14/36
4 或 5	15/36
6	17/36
7 或 8	6/36
9	5/36
10 或 12	3/36
11	2/36
15、16、18、20 或 24	1/36

77

如果对方的一枚孤子在你的某一枚棋子之前 5 格，你用这枚棋子俘虏对方的孤子的概率有多大？只要你掷出的两枚骰子

① 根据双陆棋的规则，一方的一枚棋子单独居于某一个格称为“孤子”。如果对方的某一个棋子恰好落在你的孤子所在的格，则你的孤子被对方“俘虏”。如果某一方的两个或两个以上棋子同时居于某一格，则这一格已被这一方占领，对方不可在此格落子。被俘虏的棋子必须放到棋盘的中线上，如果你想让被俘虏的棋子重新出发，必须把棋子放到对方的大本营内未被对方占领的格内，其位置由你掷出的一枚骰子的号码决定。——译者注。

中有一个结果是“5”，你就可以达到目的。有11种情况符合这个条件（1-5，2-5，3-5，4-5，5-5，6-5，5-1，5-2，5-3，5-4，5-6）。此外，如果你掷出的两枚骰子的号码之和为5，你也可以达到目的。有4种情况符合这个条件（4-1，3-2，2-3，1-4）。所以，你俘虏对方棋子的概率是15/36。如果对方的孤子距离你的棋子8格，则你俘虏对方棋子的概率要小得多，只有当你掷出的两枚骰子的号码总和是8时才行，对应的情况有6种（2-2，4-4，6-2，5-3，3-5，2-6）。所以，你俘虏对方棋子的概率是6/36。表5.2列出了各种距离对应的俘虏对方棋子的概率。

表5.3：对手在其大本营内占领的格数对应的
你为解救被俘虏棋子所需的掷骰子的次数

对手在其大本营内占领的格数	1	2	3	4	5
你想放入1枚被俘虏棋子所需的平均次数	1.03	1.25	1.33	1.8	3.29
你想放入2枚被俘虏棋子所需的平均次数	1.32	1.69	2.22	3.24	6.25

请注意当距离由6变成7时，棋子被俘虏的概率剧烈下降。当距离不超过6时，被俘虏的概率很高。如果你必须使你的孤子距敌方不超过6格，你应当使距离尽量小。不过有时候情况恰好相反，你可能故意希望对方俘虏你的棋子，这样你就可以把棋子放到对方的大本营内去干扰对手。这时你需要知道平均需要掷多少次骰子才能把你被俘虏的棋子放入对方的大本营，这个平均次数与对手在其大本营内占领的格数和你想放入棋盘的被俘虏的棋子个数有关。细节见表5.3。

假设你有1枚棋子被俘虏，而对手控制着4个格。你不能

把棋子放入棋盘的条件是两枚骰子的号码都指向对手控制的格。一枚骰子的号码指向对手控制的格的概率是4/6，两枚骰子的号码都指向对手控制的格的概率是（4/6）×（4/6）=4/9，所以，你可以把棋子放入棋盘的概率是1-4/9=5/9。也就是说，为了把1枚被俘虏的棋子放入棋盘，平均需要9/5=1.8次。附录3介绍了计算的细节。

假设你有2枚棋子被俘虏，而对手控制着4个格。你有可能只用1次就把2枚棋子都放入棋盘，条件是两枚骰子的号码都不指向对手控制的格。这个概率是（2/6）×（2/6）=1/9。此外，有4/9的可能只能放入1枚棋子，有4/9的可能1枚棋子也放不进去。综合各种情况，为了把2枚被俘虏的棋子放入棋盘，平均需要81/25=3.24次。

在实战中利用表5.2需要仔细地考虑，因为你必须考虑另一个问题：把对方的棋子俘虏可以给对手制造多大的麻烦。比如这个例子：你在自己的大本营中控制了4个格，你可以俘虏 78 对手的两枚棋子，代价是在自己的五号位为自己制造一枚孤子。此时俘虏对手的棋子很可能是上算的（依赖于其他棋子的站位）：对手为了解放自己的棋子很可能耗费数轮，而你的孤子有2/3强的概率在被俘虏之前逃脱。

在双陆棋中经常需要对概率的计算和判断。优秀的双陆棋教科书总是既强调数学运算，也强调实战经验。目前，最优秀的玩双陆棋的电脑程序已经可以在1秒以内判断数千个局面，它可以检查所有可能的局面从中选择对自己最有利的一种。国际象棋世界冠军（卡斯帕罗夫）已经被电脑击败，即使现在电脑尚未在双陆棋上征服人脑，这个日子也为期不远。

两个骰子的游戏——垄断棋

垄断棋起源于美国，在世界各地有很多种版本。为了避免版本不同造成的混乱，以下用英国垄断棋中的地产颜色表示各种地产。[①]当游戏者达到四人或更多时，一个人很难控制某一类地产的全部，所以你需要和别的玩家交换地产。此时，你需要了解各类地产的价值。

有四个因素决定某一块地产的价值：

○ 买进此地时的花费；

○ 在土地上盖建筑的投资；

○ 你可以向对手收取的租金；

○ 在棋盘上的位置。

---

① 1990年前后垄断棋在中国流行一时，国人称之为“强手棋”，是一种模拟经营地产的游戏。游戏可以有二至八人参加，每人在游戏之初持有相同数量的“货币”作为经营资金。每人有一枚棋子，轮流在棋盘上行棋，走多少步由掷骰子决定。棋盘上有“街道”，棋子在街道上行进。街道边是地产。当在无主地产边停留时，你可以投资买下这块地产。如果你在自己的地产边停留，可以决定盖某种建筑，建筑的种类受地产的性质的限制；如果你在别人的地产边停留，需要向地主支付租金，租金额取决于建筑的种类和级别以及所处地段。有些地点比较特殊：“监狱”会囚禁你若干轮，另外一些地点会把你直接送到其它地点，有时由抽一张卡片决定。一旦你拥有某一类地产的全部（所谓的“垄断”），你可以使这类地产“升级”，升级以后的地产租金猛增。游戏的目的是尽量保证自己的财务正常运转，同时迫使对手“破产”。破产后的对手出局，剩下的最后一个玩家是胜利者。台湾人在垄断棋的基础上开发了一种电脑游戏，名为“大富翁”。这种游戏已发展了多种版本，至今仍在年轻人中流行。不过对棋盘上的强手棋中国人似乎已失去兴趣。请看图5.5。——译者注。

前两个因素的意义是显而易见的。第三个因素的意义在于衡量你可以用这块地给对手以多大的打击。（豪华地段上的深蓝色建筑可以一次收取 2 000 英镑以上的租金，通常足以构成致命一击；而在偏僻地段上的棕色建筑的租金不超过 250 元。）最后一个因素的重要性通常被玩家忽视。



图 5.5:垄断棋的棋盘

棋盘上有 40 个可停留的格。如果玩家仅仅根据骰子的结果在各个格之间移动，则每个格被访问的频率是相同的。但事

实上，有些格比较特殊，在这些格上你会直接跳到另一个格，其中某些格直接把你送进监狱。监狱被访问的频率比其他格高 79 得多。在游戏中应如何利用这个特点呢？

每次棋子走的步数由两枚骰子决定，范围是从 2 格到 12 格。中间的数出现的概率较高。在所有可能中，7 出现的概率是 6/36，6 和 8 出现的概率是 5/36，5 和 9 出现的概率是 4/36，等等，两端的数出现的概率依次递减，而 2 和 12 出现的概率仅为 1/36。从监狱所在的格出发，可能立刻到达 3 个桔黄色的格，距离分别是 6 格、8 格和 9 格。因此，离开监狱直接进入桔黄色格的概率高达 14/36。此外，如果第一次掷出的结果低于 10，还有第二次进入桔黄色格的机会。相比之下，监狱附近的紫色格与监狱的距离分别是 1 格、3 格和 4 格，从监狱出发直接进入紫色格的概率仅为 1/7。

如果连续三轮掷出的骰子都是两枚骰子号码相同，你立即进入监狱。此外，在许多需要抽卡片的格上，卡片也可以直接把你送进监狱。用计算机模拟游戏的进程可以发现各个格被访 80 问的频率。计算机模拟的结果显示，当各个格被访问的总次数达到 1 000 万时，监狱被访问的频率是普通格的二倍。其它格被访问的频率也可以统计出来。

一共有六类包括三种建筑的地产。其中紫色或浅蓝色格子每被访问 100 次，绿色或黄色格被访问 110 次，桔黄色或红色格被访问 122 次。另外两种地产（深蓝色和棕色）仅包括两种建筑，因而被访问的频率低很多。为精确比较，我们统计出当平均每个紫色格被访问 100 次时，平均每个棕色或深蓝色格被访问了多少次。结果是 102。

下面看看租金与建筑数量的关系。当建筑数量达到 2 或 3 时，租金增加了很多。尤其是建筑数量达到 3 时，即使你用贷

款建房，扣除贷款利息以后租金的增长仍然非常可观。当对手不得不向你的三个建筑交租金时，这笔钱通常会让他很为难。因此，如果你有两类建筑，应当尽量使每一类建筑都有 3 个，避免数量分散。

每一种地产有独特的经济性能。对于每一种地产，我们计算出把这类地产全买下来并盖成旅馆所需的花费，称之为“投资额”；同时，计算出这类地产能获得的最高租金的总和，称之为“产出额”。算出产出额与投资额的比值，称之为“收益系数”。显然，收益系数越高对其所有者越有利。在游戏中收益系数没有明显的意义，仅提供一个比较不同的地产的标准。八种地产按收益系数的高低排序如下：浅蓝色（1.59），桔黄色（1.41），深蓝色（1.27），紫色（1.24），黄色（1.15），棕色（1.13），红色（1.09），绿色（1.01）。

除收益系数以外，另外一个需要考虑的因素是在一种地产的租金达到较高程度以前需要投入多少钱。我们以租金超过 750 英镑为标准。（750 这个数的选择相当随意，我想举出一个足以让对手感到财政困难的数额。）棕色和浅蓝色地产的租金达不到这个程度。其它六种地产按所需资金由低到高排序：桔黄色（1 760 英镑），紫色（1 940 英镑），深蓝色（1 950 英镑），黄色（2 150 英镑），红色（2 330 英镑），绿色（2 720 英镑）。深蓝色地产的实际价值要比这个顺序的显示结果高，因为它的租金比 750 英镑高很多。基于类似原因，绿色地产的实际价值也要比这个顺序的显示结果高。

81

理想的投资目标应符合以下条件：被访问的频率较高，价格合理，容易给对手制造麻烦。权衡这些因素的结果是棕色和绿色的地产很少吸引投资者的兴趣。棕色格被访问的频率较低，而绿色地产价格昂贵。相比之下桔黄色地产最受欢迎：被

访问的频率较高，价钱便宜，租金可观。另外的五种地产差别不大。当你准备和别人交换地产时，需要考虑这些因素。最后一个需要注意的要点是尽量使你的建筑多。棋盘上可以容纳的建筑总数是固定的，你占据的空间越多，留给对手的空间就越小。

垄断棋主要还是凭运气取胜。如果你掌握了什么获胜的诀窍，对手可以很快学会，并用来对付你。

两个骰子的游戏——双骰赌博

双骰赌博有许多变种，根据下注方式的不同有很多种玩法。如许多其它的骰子游戏一样，双骰赌博的历史相当悠久，但现代规则的确立非常晚。第一次世界大战中的士兵确定了双骰赌博的现代规则：你掷两枚骰子，算出两枚骰子的号码总和。如果号码总和是 7 或 11，你立刻获胜；如果号码总和是 2、3 或 12，你立刻失败；如果不是以上五种情况，则连续掷骰子。如果你最初掷出的号码总和再次出现，你获胜；如果在此之前你掷出了 7，你失败。

显然，你首轮获胜的概率是 $6/36+2/36=8/36$。如果首轮未分出胜负，则胜负取决于两件事——“最初掷出的号码总和再次出现”以及“掷出的号码总和为 7”——中的哪一个先出现。初看起来问题很复杂，因为我们不知道需要掷多少次骰子。其实问题可以大大简化：如果两个事件 A 和 B 是相互排斥的（即不能同时发生），两个事件发生的概率分别是 a 和 b，则事件 A 先于事件 B 发生的概率是 $a/(a+b)$。例如，A 表示“掷出的号码总和是 6”，B 表示“掷出的号码总和是 7”。两个

事件发生的概率分别是5/36和6/36，所以A先发生的概率是5/11。

根据在于：如果在某一轮中A和B都不发生，我们可以忽略这一轮。只有A或B发生的轮次与我们的问题有关。由于A和B发生的概率分别是a和b，在连续掷N次骰子以后（N足够大），A和B发生的次数分别是Na和Nb。由于每一轮的结果要么是A，要么是B，而且二者的排列是随机的，所以第一个结果是A的概率是Na/（Na+Nb）=a/（a+b）。

对应于你首轮掷出的号码总和，我们可以分别算出你获胜的概率，并从而算出你获胜的总概率。例如，首轮掷出6的概率是5/36，而此后你获胜的概率是5/11，所以你首轮掷出6并最终获胜的概率是5/36×5/11。类似地你可以算出首轮掷出4、5、8、9、10的情况。此外，你还有8/36的概率首轮获胜。综合起来，你获胜的概率是244/495，约等于0.493。游戏对掷骰子的一方略微不利。

在英国赌场里，每局比赛要抽出1%的赌注交给赌场。在美国，游戏的规则有许多细微的变化。不过总的来说，这种游戏像其他赌场游戏一样，输多赢少。

三个或更多骰子的游戏

许多骰子游戏的目标是使骰子的号码总和为一个确定的数。如果你能算出总和为各个数的概率，你就可以判断游戏是否对你有利。判断的方法通常是“查数”，不过查数也是有学问的。举一个例子：如何计算三枚骰子的号码总和是9或10的概率？

可以这样计算：首先列出总和为 9 或 10 的各种组合，结果是

9=（6，2，1），（5，3，1），（5，2，2），（4，4，1），（4，3，2），（3，3，3）；

10=（6，3，1），（6，2，2），（5，4，1），（5，3，2），（4，4，2），（4，3，3）。

9 和 10 对应的组合都是六种，所以总和是 9 的概率与总和是 10 的概率相等。

如果你相信这个结果，你就上当了！类似的一个圈套：“总和为 3 的组合只有一种（1，1，1），总和为 4 的组合也只有一种（2，1，1），所以总和是 3 的概率与总和是 4 的概率相等。”

83

统计组合的方法是正确的，但结果不对。为标明错误所在，我们把三枚骰子区分开，分别涂以红色、白色和蓝色。此时问题很明显：总和是 3 的状态只有一种——每个骰子都显示 1；但总和是 4 的状态却有三种——三枚骰子中的一枚显示 2，而另外两枚显示 1。因此，总和是 4 的概率为总和是 3 的概率的三倍。同样，组合（4，4，2）对应的状态有三种——三枚骰子中的一枚显示 2，而另外两枚显示 4。当三个骰子的号码各不相同时，比如组合（5，3，1），对应的状态有六种——三枚骰子中的一枚显示 5，剩下的两枚中的一枚显示 3，最后的一枚显示 1。

回到最初的问题：计算三枚骰子掷出 9 或 10 的概率。我们算出每种组合对应的状态数，然后累加：

9：6+6+3+3+6+1=25；

10：6+3+6+6+3+3=27。

这种计算方法在历史上非常著名。大科学家伽利略除了发

明望远镜、开创动力学、宣扬日心说以外，还研究过赌博。他在一个佛罗伦萨贵族的要求下写过一本小册子，目的是解释为什么三枚骰子掷出 10 的频率高于掷出 9 的频率。表 5.4 介绍了对应于几种总和的状态数。为了求出规律值，只要把状态数除以 216 即可。如果两个数的和是 21（比如 4 和 17，10 和 11），则这两个数作为总和出现的概率相同。

结论与两枚骰子的情况非常类似。前面我们已经得出结论：掷两枚骰子，号码总和为中间值（7）的概率最大，两端的值对应的概率递减。事实上，当骰子的总数超过 3 时这个规律依然生效。当骰子的数目为偶数时，中间值是惟一的；当骰子的数目为奇数时，中间值有两个，这两个值对应的概率相同。在数学上有一个公式，在给定骰子的数目和号码总和的情况下可以立刻求出对应的概率值。有兴趣的读者可以参阅理查德·艾斯坦的著作《赌博原理与统计方法》。

表 5.4：对应于几种总和的状态数

总和	3	4	5	6	7	8	9	10
对应的状态数	1	3	6	10	15	21	25	27

84

四个骰子——凯夫勒·德·米勒问题

关于概率的系统研究始于 17 世纪，开创者是两个伟大的数学家：帕斯卡和费尔马。凯夫勒·德·米勒是一个著名的赌徒，他考虑过一个问题：抛四枚骰子，至少出现一个 6 的概率和一个 6 也不出现的概率哪个大？这个问题受到帕斯卡和费尔马的重视，概率论研究由此发端。

附录 2 介绍了这个问题的答案：一个 6 也不出现的概率是 625/1 296，至少出现一个 6 的概率是 671/1 296。后者的概率较大。凯夫勒·德·米勒热衷于一个游戏：每次掷两枚骰子，目的是至少得到一次两个骰子都是 6 的情况。需要掷多少次骰子才能使至少得到一次双 6 的概率大于 1/2?

有一种解法如下：掷一枚骰子有 6 种可能结果，掷两枚骰子有 36 种可能结果；由于掷一枚骰子需要 4 次才能使得到至少一个 6 的概率高于 1/2，而 36 恰好是 6×6，所以需要掷 24 次骰子才能使至少得到一次双 6 的概率大于 1/2（24＝4×6）。

这个答案是错误的！在求掷四枚骰子至少出现一个 6 的概率时，正确的方法是先计算没有 6 出现的概率。我们用同样的方法计算每次掷两枚骰子、连续掷 24 次的情况。两枚骰子可能出现 36 种结果，出现双 6 的概率是 1/36，不出现双 6 的概率是 35/36。所以，连续抛 24 次，每次都不出现双 6 的概率是 $(35/36)^{24}$，这个数略大于二分之一（约等于 0.508 6）。此时至少得到一次双 6 的概率低于二分之一。利用同样的方法可以算出，连续掷 25 次骰子，每次双 6 都不出现的概率刚好低于二分之一（0.494 5），所以我们要找的答案是 25，而非 24。

后人通常认为，德·米勒本人也是一个数学家，他知道解出 24 的计算过程是错误的。由于掷 24 次和掷 25 次至少得到一次双 6 的概率都非常接近于二分之一，所以德·米勒通过赌博实战发现 24 不是正确答案的可能性不大。当然，有些书上记载的就是德·米勒通过实战发现了问题。这个故事再次证明，概率论是一个极容易让人犯错误的领域，而且计算上的错误不容易用实验发现。计算一定要小心、严格，否则你可能输钱！

85

后发制人

这个例子的规则和潘尼游戏不同，但里面的陷阱非常类似。

有四枚骰子，各个骰子上的号码与普通骰子不同：

红色骰子每个面都是 4；

蓝色骰子有两个面是 8，四个面是 2；

绿色骰子有三个面是 7，三个面是 1；

黄色骰子有四个面是 6，两个面是 0。

由于每个骰子的各面号码总和都是 24，所以其号码的平均值都是 4。任选一枚骰子，掷 N 次（N 足够大），则得到的号码总和应该大致相等，没有哪枚骰子有固定的优势。所以，选哪枚骰子无所谓。

游戏这样进行：你在四枚骰子中任选一枚，我在剩下的三枚骰子中任选一枚。两人各掷一次，号码较大者获胜。

我可以保证：无论你选哪一枚骰子，我都可以找到另一枚骰子，从而使我获胜的概率不低于 2/3。方法很简单：

○ 如果你选择蓝色骰子，则我选择红色骰子。蓝色骰子有 2/3 的可能得到 2，而红色骰子总是得到 4，所以我的胜率是 2/3；

○ 如果你选择绿色骰子，则我选择蓝色骰子。蓝色骰子有 1/3 的可能得到 8，如得到 8 则我必胜；在剩下的 2/3 的可能性中，蓝色骰子得到 2，而绿色骰子有 1/2 的可能得到 1，此时我仍然获胜。所以我的胜率是 1/3 + (2/3)(1/2) = 2/3；

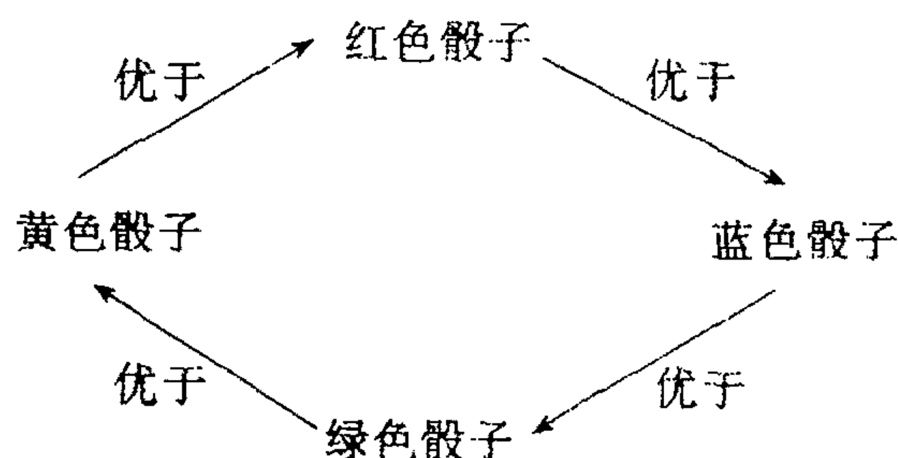
○ 如果你选择黄色骰子，则我选择绿色骰子。绿色骰子

有 1/2 的可能得到 7，如得到 7 则我必胜；在剩下的 1/2 的可能性中，绿色骰子得到 1，而黄色骰子有 1/3 的可能得到 0，此时我仍然获胜。所以我的胜率是 1/2 +（1/2）（1/3）= 2/3；

○ 如果你选择红色骰子，则我选择黄色骰子。黄色骰子有 2/3 的可能得到 6，如得到 6 则必胜。所以我的胜率是 2/3。

86

综上所述，四枚骰子的关系是：



结论可以这样表述："A 比 B 大的概率较高；B 比 C 大的概率较高；C 比 D 大的概率较高；D 比 A 大的概率较高。"但是如果你得出这个结论："红色骰子优于蓝色骰子，蓝色骰子优于绿色骰子，绿色骰子优于黄色骰子，所以，红色骰子优于黄色骰子。"那可是大错特错。如果你真的这样认为，我愿意每天都陪你赌钱！

投标游戏

投标游戏可以帮你赚钱：用粉笔在桌面上写下从 1 到 6 的号码，让参与者在六个号码上下注。下注以后你掷三枚骰子，根据被下注的号码出现的个数判断输赢。例如，我用 1 英镑赌号码 4，如果三枚骰子的结果中没有 4（比如出现 {6, 2, 3}）

则我输掉赌注；如果有一个4（比如出现{4，2，2}）则我收回2英镑（作为赌注的1英镑加上赢得的1英镑）；如果有两个4（比如出现{6，4，4}）则我收回3英镑；如果三个骰子结果都是4，则我收回4英镑。

初看起来这个游戏对庄家没什么好处。有的人可能这样考虑：每个骰子出现下注号码的概率是1/6，一共有三枚骰子，所以出现下注号码的概率是1/2，所以参与者至少有1/2的机会获胜，此外还有出现两个以上下注号码的机会，总的来说游戏对庄家不利。这种观点是错误的！正确的分析如下：每一枚骰子不出现下注号码的概率是5/6，所以三枚骰子都不出现下注号码的概率是 $(5/6)^3=125/216$，高于1/2（按照错误的思路这个概率应当等于1/2）。恰好出现一个下注号码的概率是75/216，恰好出现两个下注号码的概率是15/216，三个骰子都出现下注号码的概率是1/216。如果参与者赌1英镑，平均他将取回0.92英镑。

参与者的投资回报率是92%，游戏对庄家有利，但是这个优势还不够大。庄家赔本的危险性还是很大。为了扩大优势，聪明的庄家可以修改规则：如果结果中不出现下注号码，则下注者输掉赌注；如果结果中出现下注号码，则庄家退还赌注并赔同样数量的钱。这样，参与者的投资回报率下降到84%，平均每玩216局，庄家胜125局，负91局。 87

有一种类似的游戏用五枚扑克骰子玩。所谓“扑克骰子”与普通骰子没有根本的差别，只不过每个面上标的不是从1到6的号码，而是6个扑克牌上的号码：9，10，J，Q，K和A。游戏规则是：参与者可以任选两个面下注，下注结束以后庄家掷五枚骰子。如果下注者选的两个面都出现，则下注者获胜；如果只出现一个面或都未出现，则庄家获胜。游戏有两个极端

情况：当五枚骰子的结果完全相同时，每个下注者都失败；当五枚骰子的结果各不相同时，许多下注者获胜。

我们计算一下每个下注者的胜率。假设你选择号码 A 和 K。可能出现的情况有三种：

(1) A 和 K 在结果中都不出现；
(2) A 和 K 在结果中只有一个出现；
(3) A 和 K 在结果中都出现。

对于情况 (1)，由于一个单个的骰子结果不是 A 和 K 的概率是 $4/6=2/3$，所以情况 (1) 发生的概率是 $(2/3)^5$。

对于情况 (2)，我们先计算结果中不出现 A 但至少出现一次 K 的概率。结果中不出现 A 的概率是 $(5/6)^5$。在此前提下每个面只有 5 种可能的结果，其中 4 种不是 K，所以 K 一次也不出现的概率是 $(4/5)^5$。所以，结果中不出现 A 但至少出现一次 K 的概率是

$$(5/6)^5\times(1-(4/5)^5)$$

显然，情况 (2) 发生的概率是这个值的二倍。

由于三种情况对应的概率之和为 1，所以情况 (3) 发生的概率是

88

$$1-(2/3)^5-2\times(5/6)^5\times(1-(4/5)^5)=0.3279$$

这就是任意选择一对号码的获胜概率，低于三分之一。对于庄家而言情况极为有利。如果你来坐庄，最好能劝参与者同时买进全部的 15 种号码组合：(9，10)，(9，J)，……，(9，10)。假设他在每个组合上下 1 英镑的赌注。对于庄家而言最糟糕的情况就是掷出的五枚骰子号码各不相同，此时买家的 15 个组合中有 10 个获胜，5 个失败，你需要付给买家 20 英镑，亏损

5英镑。但是这种情况发生的概率低于十分之一，其它情况都对你有利。平均而言，你收入15英镑，支出9.84英镑。在非常罕见的情况下——概率是1/1 296——五枚骰子的结果相同，此时所有参与者的赌注都被庄家吃进。

这几个值是怎样算出来的？为了使五枚骰子的号码各不相同，第二枚骰子的号码要不同于第一枚，其概率是5/6；第三枚骰子的号码要不同于第一枚和第二枚，其概率是4/6；如此等等。所以，五枚骰子的号码各不相同的概率是（5/6）×（4/6）×（3/6）×2/6）=5/54，约等于0.092 6。这就是你输5英镑的概率值。对于买家的每一个号码组合，你支出2英镑的概率是0.327 9。在你收到买家的15英镑以后，你的平均支出是15×2×0.327 9=9.84英镑。相当于每收入1英镑支出0.66英镑。

当然骰子的个数不一定是5。需要注意的是，骰子的个数越多，庄家的风险越大。当骰子的个数是6时，游戏仍然对庄家有利——平均每收入1英镑，支出0.84英镑。当骰子个数达到7或更多时，游戏对庄家不利。 89

第六章　抉择较少的游戏

我们从一个硬币游戏开始讨论。我和你对赌，每个人选择硬币的一个面。选择独立做出，在选择过程中不要告诉对方自己的选择。选择“正面”还是“反面”可以是任意的，也可以利用抛硬币的方法随机指定：我抛一次硬币，抛出哪个面就选择哪个面；你也抛一次硬币，抛出哪个面就选择哪个面。选择过程完成以后，胜负取决于我们的选择结果：如果我们选择的结果相同，你赢得我的 2 英镑；如果我们选择的结果不同，我赢得你的 2 英镑。这个游戏是公平的，还是对某一方有利?

表 6.1 表示游戏的所有可能结果。第一列是我的选择，第一行是你的选择。+2 表示“我赢得你的两英镑”，-2 表示“你赢得我的两英镑”。

显然，这个游戏是公平的，游戏中没有陷阱。如果你不关心我选择什么，只是完全随机地做出自己的选择，则你有 1/2 的机会获胜，同样有 1/2 的机会失败。其实这个游戏和硬币没什么关系，倒是和“石头、剪子、布”类似。

下面我们改一下游戏规则。如表 6.2 所示，如果我们的选择都是正面，则我输 3 英镑；如果我们的选择都是反面，则我

输1英镑；如果我们的选择不同，则你输2英镑。此时游戏是否公平？你也许会这样想："这个游戏不是公平的。因为当你选择反面时结果对你有利，你每次都会选择反面，这样我每次只能赢1英镑，你每次却能赢2英镑。"这个想法有点傻。

表6.1：游戏的所有可能结果

	你选择正面	你选择反面
我选择正面	-2	+2
我选择反面	+2	-2

90

表6.2：改变规则之后的所有可能结果

	你选择正面	你选择反面
我选择正面	-3	+2
我选择反面	+2	-1

假如我每次都选择反面的话，你可以每次都选择反面，这样每一轮我都输1英镑，我赢2英镑的情况永远也不会发生。

但你的结论是正确的：这个游戏确实对我有利。我有办法保证平均每一轮获利1/8英镑。待会儿我会告诉你我的诀窍，现在我们先研究一个类似的游戏。这个游戏的特点是玩家的目的不是直接从对方手中赢钱，而是尽量使自己的利益最大化。

约会游戏

一个男孩和一个女孩约会，两人约好8点钟在俱乐部门口

见面。如果两个人都早到或都晚到，则他们可以度过一个美好的夜晚。此时，我们认为男孩的得分是 +1；如果男孩早到而女孩晚到，则男孩会遇到尴尬的情况——遭遇自己的前女友。此时，我们认为男孩的得分是 -1；最后，如果女孩早到而男孩晚到，女孩会遇到自己的前男友，把男孩甩在一边和前男友走。对于男孩而言这是最糟糕的情况，此时我们认为男孩的得分是 -2。男孩的得分情况如下表（数学家称之为"支付矩阵"），他应当如何决定？

	女孩早到	女孩晚到
男孩早到	1	-1
男孩晚到	-2	1

在男孩和女孩之间没有赌注的流通，但两个人都希望自己今晚的快乐最大化（也就是得分最大化，应为得分即模拟快乐程度）。女孩也要决定早到还是晚到，她用自己的标准判断四种可能情况给自己带来的快乐。她会画出自己的"表"（即支付矩阵）。女孩的表可能与男孩的表完全不同，不过男孩并不关心这种差别，他只是按照自己的表做决定。你可能认为男孩的态度是自私的，但是男孩的目的不过是度过一个快乐的夜晚，没什么可以指责的。同样，女孩也不关心男孩的表与自己的差异，仅仅按自己的表行事，判断哪一种选择可以使自己的欢乐最大化。如果两个人的选择不合拍，他们很快就能发现。

91

约会游戏与表 6.2 表示的游戏类似，可以用相同的方法解决。但是我们首先要分析一类相似的游戏，这些游戏比较容易分析，而且可以显示我们讨论的问题的独特之处。

一个简单的游戏

这个游戏在两个人——罗伊和克林之间进行。罗伊有两种选择：a 和 b，克林也有两种选择：A 和 B。在两个人都确定自己的选择之后，根据两个人的选择判断游戏的结果。

罗伊赢得的就是克林输掉的，罗伊只关心使自己的收益最大化。双方在利益上的冲突与表 6.2 表示的情况一样。这与约会游戏略有不同，在约会游戏中男孩的所得与女孩的所失无关。表 6.3 表示在各种情况下罗伊的收益（罗伊的支付矩阵）。从罗伊的角度看，他担心克林会洞悉他的策略，所以要考虑克林对他的估计。比较 a 和 b：如果选择 a，保证最少赢 2；如果选择 b，只能保证最少赢 1。所以对罗伊来说，选择 a 优于选择 b。另外，从克林的角度看，克林也担心罗伊会洞悉他的策略。比较 A 和 B，选择 A 最多可能输 4，而选择 B 最大可能输 2。所以对克林来说，选择 B 优于选择 A。

表 6.3：在各种情况下罗伊的收益

	克林选择 A	克林选择 B
罗伊选择 a	4	2
罗伊选择 b	3	1

92

现在游戏有一个确定的结果：罗伊选择 a，而克林选择 B。两个人谁也不敢改变自己的选择。如果罗伊把自己的选择改为 b，则他的收益从 2 减少到 1；同样，如果克林把自己的选择改为 A，则他的支出从 2 增加到 4。两个人的利益均衡产生了

一个稳定点：(a，B)。

然而在表 6.2 中没有这样的稳定点。在表 6.2 中，如果我选择正面，我的最坏结果是 -3；如果我选择反面，我的最坏结果是 -1。所以我选择反面优于选择正面。而从你的角度看，选择正面或反面没有明显的优劣。在这里，不存在一个确定的结果使我们两个谁也不敢改变选择。在约会游戏中情况也是如此。这两个例子比表 6.3 更复杂。

稳定点是这样确定的：

(1) 找出每一行中的最小值，然后找出所有最小值中最大的一个。这个值称为“最大极小值”；

(2) 找出每一列中的最大值，然后找出所有最大值中最小的一个。这个值称为“最小极大值”。

如果“最大极小值”和“最小极大值”重合，则稳定点存在。在多数情况下这两个值不重合，但是每当我们面临一个新问题时，我们首先应当考虑这两个值是否重合，这样可以使我们少走很多弯路。当每个参与者可以选择的策略多于两种时这个方法依然有效。

钓鱼问题

罗伊和克林两人相约去钓鱼。他们约定在 12 个湖中选择一个作为目的地。这 12 个湖位于三条东西方向的公路和四条南北方向的公路的交叉点上，确定地点的方法是罗伊在三条东西方向的公路中选择一条，克林在四条南北方向的公路中选择

一条，其交叉点即为目的地。每个湖中只有两种鱼——梭子鱼和鲑鱼。罗伊喜欢钓梭子鱼，想去梭子鱼密度高的湖；克林恰好相反，喜欢钓鲑鱼。由于梭子鱼的密度高就意味着鲑鱼的密度低，所以克林想去梭子鱼密度低的湖。表 6.4 列出了各个湖中梭子鱼的密度，a、b、c 表示东西方向的公路，A、B、C、D 表示南北方向的公路。

罗伊发现三条公路 a、b、c 对应的湖中的最低梭子鱼密度分别是 1、4、2，其中的最大值是 4，所以如果他选择 b，可以保证梭子鱼密度不低于 4。另外，克林发现四条公路 A、B、C、D 对应的湖中的最高梭子鱼密度分别是 5、7、4、8，其中的最小值是 4，所以如果他选择 C，可以保证梭子鱼密度不高于 4。此时，“最大极小值”和“最小极大值”重合于（b，C），这就是稳定点。

93

表 6.4：12 个湖中梭子鱼的密度（个数/公顷）

	A	B	C	D
A	3	7	2	1
B	5	5	4	6
C	4	2	3	8

一旦他们认识到这个道理，就无需为自己的选择保密。罗伊可以坦白地告诉克林自己将选择 b，克林即使知道了这个消息，最佳选择也只能是 C。同样，克林可以坦白地告诉罗伊自己将选择 C，克林即使知道了这个消息，最佳选择也只能是 b。

如果表 6.4 中的数字表示的不是梭子鱼的密度，而是湖的海拔高度（以百米单位），则（b，C）决定的湖是地势中的鞍点。从南北方向看，这个湖的地势最高；从东西方向看，这个

湖的地势最低。[1] 解决问题的过程通常是找到“最大极小值”和“最小极大值”，看看这两个值是否重合，如果重合则为鞍点。这个过程可以称之为“寻找鞍点”。

本章的目的是讨论抉择较少的游戏，所以我们不会设计表中的数据非常多的情况。钓鱼问题的经验是在双方有多种选择时首先检验鞍点是否存在。当双方的选择较少时也是如此。

回到表 6.2

表 6.2 中没有鞍点，但是我有办法保证平均每局赢 1/8 英镑。现在用罗伊和克林分别替代我和你。在没有鞍点的游戏中，一个玩家不能始终坚持同一个选择，而应当随机地做出决定。此时，最关键的问题在于确定每一个选择对应的频率。在表 6.2 中，罗伊的最佳策略是选择反面多一些。确切地说，在每一轮中罗伊应当以 3/8 的概率选择正面，以 5/8 的概率选择反面，并保证每一轮的选择相互独立。这种选择称为“混合策略”。

下面我们来证明这个策略的优越性。假设克林选择正面。此时我们把注意力集中于表 6.2 中的第一列数字。如果罗伊选择正面（概率为 3/8），罗伊“赢” －3 英镑；如果罗伊选择反面（概率为 5/8），罗伊赢 2 英镑；所以在此情况下罗伊平均赢

$$(-3)\times(3/8)+(2)\times(5/8)=1/8$$

[1] 鞍点是对策论中的常用词汇。之所以称为“鞍点”，是因为马鞍的中心点具有同样的性质。——译者注

假设克林选择反面。此时我们把注意力集中于表 6.2 中的第二列数字。如果罗伊选择正面（概率为 3/8），罗伊赢 2 英镑；如果罗伊选择反面（概率为 5/8），罗伊“赢” -1 英镑；所以在此情况下罗伊平均赢

$$(2) \times (3/8) + (-1) \times (5/8) = 1/8$$

这说明，无论克林如何选择，罗伊都将平均赢 1/8 英镑。和钓鱼问题中的情况一样，罗伊甚至无需为自己的决策保密。即使对手知道了罗伊的策略，也无法阻止罗伊在游戏中获利。平均每玩 8 轮，罗伊赢 1 英镑；平均每玩 80 轮，罗伊赢 10 英镑。罗伊不能保证必胜，但在游戏中保持优势。

如果罗伊已经知道了他的最佳策略是以 3/8 的概率选择正面，以 5/8 的概率选择反面，他如何才能保证自己的每次抉择既是随机的，又符合预期的概率？我介绍两个办法。这两个办法都需要器械的帮助。第一个办法需要一只电子秒表，读数精确到百分之一秒。在每次决定之前，开启秒表，过一会再让秒表停止，看看秒表读数的倒数第二位数字是几。这个数字显然是 {0, 1, 2, …, 9} 中的一个，而且是其中任何一个的概率相等。如果是 0 或 9，这次不算，再让秒表跑一次，直到得到的数字不是 0 和 9 为止。如果得到的数字是 {1, 2, 3} 中的一个，则这一次你选择正面；否则，选择反面。这个方法可以保证每次选择是随机的，而且选择正面的概率和选择反面的概率符合要求。

第二个办法要借助于三枚硬币。我们知道，抛三枚硬币可能有八种结果，每种结果出现的概率相同，而且其中正面恰好出现一个的情况有三种。所以罗伊在决定之前可以抛三枚硬币，如果结果中恰好出现一个正面，则选择正面；否则选择反面。需要注意的是，借助于器械做出随机的决定比依赖于自己

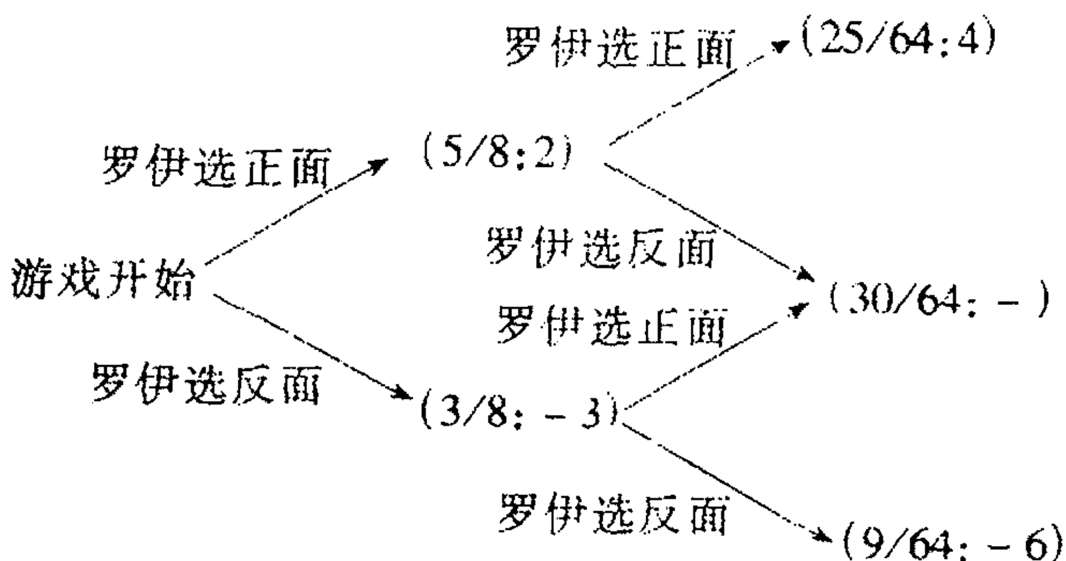
的大脑可靠得多！

如果你想在这类游戏中获胜，必须装备生成随机数字的器械。电子秒表是很好的选择，既方便，又精确。此外，你可以选择硬币、骰子、扑克牌等等，事先需要一些计算。

在这个游戏中，平均每轮罗伊可以获利 1/8 英镑。这个优势非常微妙，对手通常需要很长时间才能识破机关。相反，在潘尼游戏（详见第四章）中，后选择的一方优势非常明显，其胜率至少比对手高一倍，对手很快就会发觉形势不对。而在表 6.2 的例子里，如果对手随机地选择正面和反面，并保持选择正面和反面的频率相等，则双方获胜的次数基本上同样多，只不过罗伊输 1 英镑的次数比输 3 英镑的次数多得多，对手却每次都输 2 英镑。这个诀窍实在不容易被察觉。

假如罗伊按这个策略玩了若干局，他在游戏中获利的概率是多少？这个问题依赖于两个因数：对手的策略以及一共进行多少局。我们首先考虑最简单的一种情况：对手每次都选择正面。此时每局只有两种可能结果：罗伊赢 2 英镑或罗伊输 3 英镑。图 6.1 显示了游戏进行两局以后的结果。

图 6.1：游戏进行两局以后罗伊的盈利状况及对应的概率



表中箭头上的文字表示罗伊在本局中的选择，箭头之后的文字表示此选择的概率以及在此选择之下罗伊的盈利状况。结论令人吃惊：当两局结束之后，虽然从整体上说罗伊有利可图（平均收入 1/4 英镑），但是仅在 25/64 的情况下罗伊的盈利值是正数，而在多数情况下（概率为 39/64）罗伊的盈利值是负数。即罗伊输钱的概率比赢钱的概率高。

如果对手随机地选择正面或反面，则问题复杂得多。每一局游戏有三种可能结果：罗伊赢 2 英镑（概率为 1/2），罗伊输 1 英镑（概率为 5/16）以及罗伊输 1 英镑（概率为 3/16）。在两局之后，结论有了变化：罗伊不仅在整体收入上占优势，其盈利的机会也优于对手。

在对手选择正面和反面的频率相等的情况下，游戏进行很多局以后的情况如何？游戏进行 50 局或 100 局以后，罗伊盈利的概率是多少？为了保证罗伊盈利的概率高于 90%，游戏至少应进行多少局？回答这些问题并不困难，但计算过程非常非常繁琐。这类问题最好用计算机解决。

表 6.5 是根据计算机模拟游戏过程的结果统计出的数据。为了保证数据的可靠性，每一个数据都是根据 100 万次以上的计算机模拟统计出来的。在游戏进行 2、4、8，……局以后，罗伊的平均收益分别是 2、4、8，……英镑。但是，即使游戏进行了很多局，罗伊失利的可能性还是存在的。游戏进行 500 局以后，罗伊还有 8%的概率失利。

表 6.5：游戏进行若干局之后的战况（计算机模拟的结果）

游戏进行的局数	16	32	64	128	256	512	1024
罗伊盈利的概率（%）	58	62.5	69	76.6	84.5	92	97
双方收益持平的概率（%）	5	3.3	2	1.2	0.7	0.3	0.1

如何使游戏公平？

在约会游戏中，无所谓公平与否的问题，因为在两个参与者之间没有赌注的流通。但是在参与者通过游戏决定赌注的归属的情况下，公平与否是必须考虑的。所谓公平，是指每个参与者在游戏中的平均收益都等于 0。在表 6.1 表示的游戏中，游戏本身就是公平的。而在表 6.2 中，游戏是不公平的。为了使一个这样的游戏变成公平的游戏，应当补充一条规定：在每一局开始之前，占优势的一方应当支付给对手一定数量的筹码。

在表 6.2 表示的游戏中，罗伊可以保证平均每局获利 1/8 英镑。这也是他可以保证的最高收益，选择任何策略都不可能使他的平均收益更高。在这个游戏中，如果我们规定每一局之前罗伊付给对手 1/8 英镑，则游戏就是公平的。

97

对于存在鞍点的游戏，使其变成公平游戏非常简单。占优势的一方需要付给对手的钱就是鞍点位置上的数值。在钓鱼问题中，双方的收益不是用货币的形式表达的，所以我们先要把相关数据转换成用钱表示的形式。在一个湖中，梭子鱼的密度越高，则在此湖钓鱼对罗伊越有利，对克林越不利。罗伊和克林应当达成一个协议，依据每个湖中梭子鱼的密度对事件“去

此湖钓鱼”赋值，这个值表示罗伊应当付给克林的钱数。在（a，D）、（c，B）之类的湖中，梭子鱼的密度较低，对克林越有利，所以对应的值应当是负数；相反，在（c，D）之类的湖中，梭子鱼的密度较高，对罗伊越有利，所以对应的值应当是正数。无论罗伊和克林如何约定，（c，D）对应的值都是鞍点。如果罗伊付给克林相当于鞍点数值的钱，游戏就是公平的。

如何发现最佳策略

对于不存在鞍点的游戏，如果双方的选择都只有两种，则最佳策略很容易发现。我们用表 6.2A 为例说明发现过程。表 6.2A 是根据表 6.2 变化而来的，把表 6.2 中的每个数值都加上 10 就得到表 6.2A。这样做的目的是消除表中的负数，而这个变化对最佳策略没有影响。首先我们求罗伊的最佳策略。

表 6.2A：表 6.2 的变种

	克林选择正面	克林选择反面
罗伊选择正面	7	12
罗伊选择反面	12	9

1. 用第一列数减去第二列数，得到：

-5
+3

2. 删去正负号，并交换两个数的位置，得到：

3
5

3. 求出两个数的和。在本例中，3 + 5 = 8。罗伊选择正面和反面的频率应分别是 3/8 和 5/8。

克林可以用类似方法算出自己的最佳策略：

1. 用第一行数减去第二行数，得到：

-5	3

2. 删去正负号，并交换两个数的位置，得到：

3	5

3. 求出两个数的和。在本例中，3 + 5 = 8。罗伊选择正面 和反面的频率应分别是 3/8 和 5/8。

98

在本例中两个人的最佳策略相同，不过这是特殊情况。在多数情况下对手之间的最佳策略不同。

表 6.2A 是通过把表 6.2 中的每个数值都加上 10 得到的，而对求最佳策略的过程略加推敲即可发现，每个数值都加 10 对计算结果没有任何影响。所以，表 6.2 中的最佳策略与表 6.2A 完全相同。请读者验证这个结论。为便于阅读，我们把表 6.2 列在这里：

表 6.2

	你选择正面	你选择反面
我选择正面	-3	+2
我选择反面	+2	-1

用表 6.2 中的第一列数减去第二列数，得到 -5 和 +3；删去正负号并交换位置得到 3 和 5；结论是分别以 3/8 和 5/8

的概率选择正面和反面。

表 6.6 留给读者做练习，这个表也是表 6.2 的变种。正确答案是：罗伊以 1/4 的频率选择正面，而克林选择正面和选择反面的频率相等。

在求得最佳策略之后，每个参与者的平均收益很容易计算。具体方法是把第一列中的两个数字分别乘以相应的频率，然后求和。例如在表 6.2 中，第一列中的两个数（-3 和 +2）对应的频率分别是 3/8 和 5/8，所以罗伊的平均收益是

(-3) × (3/8) + (2) × (5/8) = 1/8

这个值也可以利用第二列中的数值和对应的频率来计算，计算方法相同，结果也一样。我们可以利用这个性质来验证求得的最佳策略是否正确。在表 6.6 中，罗伊的平均收益是

(-3) × (1/4) + (1) × (3/4) = 0

这说明表 6.6 对应的游戏是公平的。

表 6.6

	克林选择正面	克林选择反面
罗伊选择正面	-3	3
罗伊选择反面	1	-1

我们把罗伊的平均收益称为游戏的“公平参数”。如果一个游戏不是公平的，只要罗伊在每局之前付给对手相当于公平参数的钱，游戏就变成公平的。表 6.2 中的公平参数是 1/8。由于表 6.2A 是通过把表 6.2 中的每个数值都加上 10 得到的，所以表 6.2A 中的公平参数是 10 + 1/8。所以在表 6.2A 对应的游戏中，在每局之前罗伊应当付给对手 10 + 1/8 英镑。

99

通过最佳策略的计算过程很容易发现两个规律：

1. 把表中的每个数字都加上一个相同的数，对计算结果没有影响；

2. 把表中的每个数字都乘以一个相同的数，对计算结果没有影响。

在实战中，这两条规律可以帮助你简化计算过程。

其他游戏

下面我们利用求最佳策略的方法分析几个游戏。这些游戏有两个共同点：游戏双方的利益相冲突；每个参与者希望自己的利益最大化。

约会游戏

在约会游戏中，男孩的支付矩阵如下：

	女孩早到	女孩晚到
男孩早到	1	-1
男孩晚到	-2	1

100

表中没有鞍点，所以男孩应当使用混合策略（即按一定概率随机地决定每一局中的选择）。应用计算最佳策略的方法：用第一列减去第二列，得到（2，-3）；消去正负号，交换数

值的位置，得到（3，2）；最后得出结论：男孩应当以 3/5 的概率早到，以 2/5 的概率晚到。在每一局中男孩的平均收益（表示他在当晚的快乐程度）为

$$(1)\times(3/5)+(-2)\times(2/5)=-1/5$$

结果对男孩不利。也许他应当取消约会，或者重新评价他的快乐标准。

另一方面，女孩的支付矩阵如下：

	男孩早到	男孩晚到
女孩早到	1	3
女孩晚到	2	1

（这个表仅表示一种可能性。假设女孩乐于见到男孩的尴尬，而如果与自己的前男友重逢，她会更快乐。）用第一列减去第二列，得到（-2，1）；消去正负号，交换数值的位置，得到（1，2）；最后得出结论：女孩应当以 1/3 的概率早到，以 2/3 的概率晚到。女孩的平均收益为

$$(1)\times(1/3)+(2)\times(2/3)=5/3$$

如果我们对上表的数值加以调整，把 3 换成 -1（这表示讨厌与前男友重逢），则表中出现鞍点。女孩会自动选择晚到。

有罪的丈夫

丈夫忘记了今天是不是妻子的生日。他回家的时候是否应当买一束花作为礼物送给妻子？计算最佳策略的方法在这里依然有效。

假设今天不是妻子的生日。如果丈夫没有买花，属于正常

情况，此时计丈夫的得分为0；如果丈夫买花，妻子会高兴，此时计丈夫的得分为+5。

假设今天是妻子的生日。如果丈夫没有买花，后果是灾难性的，此时计丈夫的得分为-20；如果丈夫买花，妻子会非常高兴，此时计丈夫的得分为+8。

下表是丈夫的支付矩阵。这个游戏由丈夫单独参与，没有人和他对阵。

	今天不是妻子的生日	今天是妻子的生日
不买花	0	-20
买花	5	8

101

首先检查此表是否存在鞍点。两行中的最小值分别是-20和+5，最小值中最大的是+5；两列中的最大值分别是5和8，最大值中最小的也是+5。所以鞍点存在。结论：无论今天是不是妻子的生日，丈夫都应该买花。（我认为每一个妻子都应当向自己的丈夫推荐这一段论证。）

潜艇问题

在军事行动中，敌对双方拥有的选择通常超过两个。随着每一方的选择的增加，计算最佳策略变得非常复杂。本书仅讨论最简单的情况，如果读者希望了解复杂的问题，可以参考J·D·威廉姆斯的著作《决策专家》（The Compleat Strategyst）的最后一章。本书的目的是让普通读者可以接受，所以我们只介绍每一方只有两种选择的例子。

假设你是海军司令。你指挥两艘潜艇靠近敌军的航空母舰

发动攻击。潜艇 A 携带鱼雷，潜艇 B 携带牵制武器。两艘潜艇必须穿越由敌军控制的海峡。海峡过于狭窄，两艘潜艇不能并肩行驶，必须一先一后。在通过海峡时，敌军可以进行攻击，但只能攻击一艘潜艇。如果敌军攻击第一艘潜艇，则受攻击的潜艇生存的概率是 80%；如果敌军攻击第二艘潜艇，则受攻击的潜艇生存的概率只是 60%，因为此时敌军的准备更加充分。假如携带鱼雷的潜艇成功地穿越海峡，则可以保证完成任务。为了使完成任务的概率最大，你应当如何安排潜艇穿越海峡的次序？

在这个例子中，双方各有两个选择。你的选择 a 是携带鱼雷的潜艇先行，你的选择 b 是携带鱼雷的潜艇后行；敌军的选择 A 是攻击先行的潜艇，敌军的选择 B 是攻击后行的潜艇。在各种情况下，你完成任务的概率（以百分比表示）如下表所示：

	敌军攻击先行的潜艇	敌军攻击后行的潜艇
携带鱼雷的潜艇先行	80	100
携带鱼雷的潜艇后行	100	60

在这个表中不存在鞍点，所以你应当选择混合策略。利用前文介绍的求最佳策略的方法：用第一列减去第二列，得到（－20，40）；消去正负号，交换数值的位置，得到（40，20）；最后得出结论：应当以 2/3 的概率命令携带鱼雷的潜艇先行，以 1/3 的概率命令携带鱼雷的潜艇后行。完成任务的概率是

$$(80)\times(2/3)+(100)\times(1/3)=260/3$$

所以，你有 87% 的机会完成任务。

从敌军的立场上考虑，最佳策略为以 2/3 的概率攻击先行的潜艇。当然，这要求敌军知道如何计算最佳策略。

如果一个海军指挥官不懂这个方法，他可能这样考虑：由于 80%大于 60%，所以命令携带鱼雷的潜艇先行更加有利。按照这种想法行事，他只能保证 80%的成功率。如果他了解我们讨论的方法，他就可以把成功率提高到 87%。

《破坏者》问题

戴斯芒得德·拜里的惊险小说中的某些故事涉及到我们正在讨论的概率问题。在《破坏者》的结尾，主人公需要决定向哪个方向逃跑，这个决定要用到求最佳策略的方法。这个故事中还有一个情节，主人公利用高超的数学技巧引诱别人和他赌博，而赌博对主人公有利。这个赌局我们留到第七章介绍。

下面的情节是根据拜里的故事改编的：

> 邦德被关押了很久，他已经不知道今天是星期几了。他从牢房逃出来，挣扎到码头。当他看到码头上停着一艘快艇时，他的神志清醒了。他可以向南方逃，也可以向北方逃。向哪个方向逃更安全？向南方逃距离安全地点较近，而向北方逃没有足够的燃料。布洛费尔德肯定会追他，而且布洛费尔德知道所有这些情况。邦德的思路非常敏捷，他拿出小本子开始计算，然后停下来。真糟糕！他没有手表！这时邦德发现了电台，他马上和总部的 Q 取得了联系。Q 焦急地发问：“007，你他妈的在哪？”邦德回答：“先别管我在哪。今天是星期几？”“星期一。”星期一，邦德决定向北逃。

我来解释一下邦德的决策过程。邦德有两个选择：向南方逃或向北方逃；布洛费尔德也有两个选择：向南方追或向北方追。每种情况下邦德逃脱的概率（用百分比表示）如下表：

	布洛费尔德向南方追	布洛费尔德向北方追
邦德向南方逃	70	100
邦德向北方逃	80	40

利用我们前面介绍的计算方法，邦德应当以 4/7 的概率向南方逃，而以 3/7 的概率向北方逃。由于邦德手边没有产生随机数的器械，他问 Q 今天是星期几。如果是星期日、星期一和星期二这三天中的一天，则向北方逃；否则向南方逃。如果从布洛费尔德的立场分析这个问题，布洛费尔德应当以 6/7 的概率向南方追。[①]

囚徒悖论

罗伊和克林合伙抢了一家银行。警察知道是他们两人作的案，但没有证据。警察有充分的证据证明他们两个犯有盗窃罪，但希望以抢银行的罪名向他们起诉。警察把他们两个单独关押，分别审讯，希望其中的某个人能告发其同伙。只要有一个人告发同伙，就可以按抢银行的罪名定罪，而告发者可以获

---

① 其实邦德有很多产生随机数的方法。比如说，从岸边抓起一把小石子，计算小石子的总数除以 7 得到的余数是多少。如果余数是 0、1 或 2，则向北方逃；否则向南方逃。当你需要随机数的时候，类似的方法非常多。——译者注。

得减刑。

罗伊可能这样考虑："如果我们两个都保持沉默，则只能按盗窃定罪，每人入狱2年；如果我告发克林，而克林保持沉默，则我将被当庭释放，克林入狱6年；反过来，如果克林告发我，而我保持沉默，则克林将被当庭释放，我入狱6年；如果我们两个互相告发，我们两个都入狱5年。我有两个选择：选择'合作'，则不告发克林；选择'背信'，则告发克林。我的支付矩阵如下：

	克林选择合作	克林选择背信
我选择合作	-2	-6
我选择背信	0	-5

这个表中有鞍点。结论是无论克林如何选择，我都告发他。"

克林会以完全相同的方式做出决定。所以结果两个人都选择背信，各自入狱5年。

104

根据道德的原则他们应当信任同伴并关心同伴的利益。可是他们担心，如果自己信任对方而对方出卖自己，则出现最糟糕的局面。这样互相猜忌的结果是两个人都受损失。如果两个人都信任对方，保持沉默，则结果对两个人最佳：各自入狱2年。

美女与警察局长

这个问题来自一出著名的戏剧。问题涉及到两个主要人物：美女桃斯卡和警察局长斯卡皮亚。警察局长已经下令处死

桃斯卡的情人卡瓦拉多斯，但是警察局长愿意用卡瓦拉多斯的自由和桃斯卡做交易。美女和局长达成协议：美女向局长以身相许，而局长释放美女的情人。两者都可以出卖对方。局长出卖美女的方法是派人杀掉美女的情人，而美女出卖局长的方法是用一把暗藏的匕首杀死局长。

美女的目标是营救情人和处死局长，局长的目标是拥有美女和消灭情敌。美女会这样考虑："如果局长决定不杀我的情人，我最好杀掉局长，这样就可以和情人远走高飞；如果局长决定杀我的情人，我最好杀掉局长为情人报仇。我的支付矩阵如下：

	局长信守诺言	局长出卖我
我信守诺言	5	-10
我出卖局长	10	-5

无论局长如何选择，我一定要杀死他。"

类似地局长会想："如果美女信守诺言，我最好处死她的情人，免得他碍事；如果美女决定杀我，我也应当处死她的情人，让她也不能如愿。我的支付矩阵如下：

	美女信守诺言	美女出卖我
我信守诺言	5	-10
我出卖美女	10	-5

无论美女如何选择，我一定要处死她的情人。"

这不是一个欢乐的结局。美女和局长都决定，无论对方如何行事，必然出卖对方。两个人的结局都是悲剧性的：局长失去生命，美女失去情人。这个问题中的表与囚徒悖论中的表数值不同，但意义是完全一样的。表中的数值决定参与者之间不可能形成信任关系。在所有结果中，最糟糕的情况就是自己信任对方而对方出卖自己。为了避免这种情况，必须先出卖对方。

一方有两种选择，另一方有多种选择的问题

一方有两种选择，另一方有多种选择的问题并不总是很复杂，有时候可以非常轻松地解决。假设参与者依然是罗伊和克林。面对一个问题时，首先看是否存在鞍点，如果存在，则问题已经解决。第二步，如果不存在鞍点，则检查克林的各个选择，看看是否存在一个选择明显优于另外一个选择。如果存在这种情况，则明显对克林不利的选择可以忽略不计，问题因而得到简化。由于表中的数值表示罗伊的收益，所以克林希望相应的数值尽可能地小。例如，罗伊的支付矩阵可能是这样：

	A	B	C	D
a	2	3	6	7
b	5	6	3	4

表中没有鞍点。但是比较 C 和 D 所在的列，对于克林而言，无论对手选择 a 还是 b，克林选择 C 总是比选择 D 更好。所以

D可以忽略不计，因为有理性的克林不可能选择D。同样，比较A和B所在的列，对于克林而言A总是比B更好，所以B也可以删除。这样，支付矩阵简化为：

	A	C
a	2	6
b	5	3

这个问题我们会处理。罗伊应当分别以1/3和2/3的概率选择a和b，而克林应当以相同的概率选择A和C。

如果一个选择总是比另一个选择好，则称之为“优先选择”。在上例中，对于克林而言，C是相对于D的优先选择，A是相对于B的优先选择。所以在不存在鞍点的问题中，首先应当检查是否存在优先选择。在删除所有不必要的选择以后，如果问题还原成每个人只有两种选择的情况，问题就变得非常简单。

优先选择的存在有时候和鞍点的存在相关。当双方的选择都只有两种时，鞍点存在的先决条件是至少有一个参与者有优先选择。所以当双方的选择都只有两种时，我们可以直接检查是否存在优先选择。在删除不必要选择之后，问题变得非常直观。在“有罪的丈夫”的例子中，问题就可以如此处理。对于丈夫而言，“买花”相对于“不买花”是优先选择，删除“不买花”的选择之后丈夫就只有一个选择了。在“囚徒悖论”中也是如此，对于任何一方而言“背信”总是优于“合作”。不过也有这样的情况：鞍点确实存在，但并不存在优先选择。这种情况总是出现在至少有一方的选择不只两种时。例如在“钓

106

鱼问题”中，鞍点存在而优先策略不存在。

筹码游戏

罗伊有两枚筹码，每次可以出一个筹码，也可以出两个筹码。克林有三枚筹码，每次可以出一个、两个或三个筹码。两个人各自独立地决定出几枚筹码，并注意保密。在双方都决定好之后，检验两个人的选择。游戏的胜负取决于双方出的筹码总和。如果总和是偶数，克林付给罗伊数额等于此总和的钱；如果总和是奇数，罗伊付给克林数额等于此总和的钱。

罗伊的支付矩阵如下：

	克林出 1	克林出 2	克林出 3
罗伊出 1	2	-3	4
罗伊出 2	-3	4	-5

这个表中没有鞍点，也没有优先选择。我们需要一种新方法解决这个问题。新方法的核心在于：**对于克林而言，一定存在一种最佳策略，这个最佳策略只用到两种选择。**我们首先研究如何利用这个原则发现最佳策略，然后我将介绍一种解决这类问题的几何方法。

以（1，2）表示“克林的策略由‘出 1 枚筹码’和‘出 2 枚筹码’两个选择组成”，以（1，3）表示“克林的策略由‘出 1 枚筹码’和‘出 3 枚筹码’两个选择组成”，以（2，3）表示“克林的策略由‘出 1 枚筹码’和‘出 2 枚筹码’两个选

择组成"。由于克林的最佳策略是由两个选择构成的，所以只有三种可能性。我们可以逐个检验这三种可能性。 107

首先，假设克林的最佳策略属于（1，2）。此时，克林选择"出 3 枚筹码"的可能性已被排除，问题还原为双方的选择都只有两种的情况。这个问题我们可以解决。计算出在此情况下双方的最佳策略。在罗伊选择最佳策略的情况下，计算出克林分别选择三种策略时罗伊的收益，得出的三个值中最小的一个即罗伊的平均收益。用 P_1 表示此时罗伊的平均收益。

其次，假设克林的最佳策略属于（1，3）。此时，克林选择"出 2 枚筹码"的可能性已被排除，问题还原为双方的选择都只有两种的情况。计算出在此情况下双方的最佳策略。在罗伊选择最佳策略的情况下，计算出克林分别选择三种策略时罗伊的收益，得出的三个值中最小的一个即罗伊的平均收益。用 P_2 表示此时罗伊的平均收益。

最后，假设克林的最佳策略属于（2，3）。此时，克林选择"出 1 枚筹码"的可能性已被排除，问题还原为双方的选择都只有两种的情况。计算出在此情况下双方的最佳策略。在罗伊选择最佳策略的情况下，计算出克林分别选择三种策略时罗伊的收益，得出的三个值中最小的一个即罗伊的平均收益。用 P_3 表示此时罗伊的平均收益。

比较 P_1、P_2和 P_3，其中最大的数就是罗伊在这个游戏中的平均收益，这个数对应的双方的最佳策略就是这个游戏中的双方的最佳策略。

具体的计算过程留给读者，下表是计算结果。如果你没看懂，没关系。只要你明白了下表的含义，你就知道我的意思了。

克林的选择	罗伊的最佳策略	克林出1枚筹码	克林出2枚筹码	克林出3枚筹码
(1, 2)	7/12	-1/12	-1/12	1/4
(1, 3)	1/2	-1/2	1/2	-1/2
(2, 3)	9/16	-3/16	1/16	1/16

这个表的意义比较复杂。我们以第一行数值为例解释各个数值的含义。(1, 3) 表示克林的策略中只包含两个选择："出1枚筹码"和"出2枚筹码"；7/12 表示在克林选择 (1, 2) 的情况下罗伊的最佳策略，7/12 是此时罗伊选择"出1枚筹码"的概率；-1/12 表示在罗伊选择最佳策略的情况下（即罗伊以 7/12 的概率选择"出1枚筹码"），如果克林每次都选择"出1枚筹码"，罗伊的平均收益；第二个 -1/12 表示在罗伊选择最佳策略的情况下，如果克林每次都选择"出2枚筹码"，罗伊的平均收益；1/4 表示在罗伊选择最佳策略的情况下，如果克林每次都选择"出3枚筹码"，罗伊的平均收益。在最后三个数中最小的是 -1/12，所以 $P_1 = -1/12$。

在这个问题中，$P_1 = -1/12$、$P_2 = -1/2$、$P_3 = -3/16$，显然 P_1 最大。因此，罗伊的最佳策略是以 7/12 的概率选择"出1枚筹码"。罗伊的平均收益是 -1/12，这个数也就是这个游戏的公平参数，即如果在每局之前罗伊付给克林数额相当于 -1/12的钱（负号表示事实是克林付给罗伊钱），游戏就是公平的。克林的最佳策略应当由前两种选择组成，一定不能选择"出3枚筹码"。一旦克林选择"出3枚筹码"，罗伊平均可以赢数额相当于 1/4 的钱。游戏本来是对克林有利的，这样一来克林就亏本了。

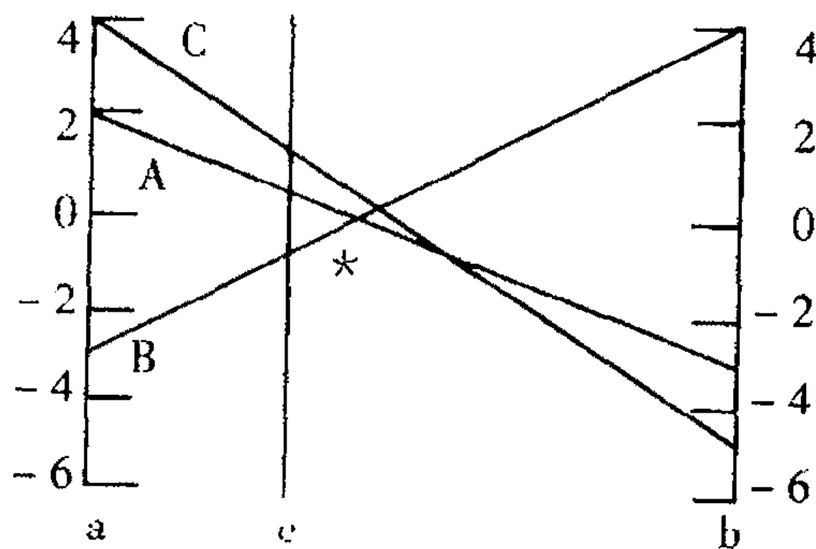
克林的最佳策略是以 1/2 的概率选择"出1枚筹码"，以

1/2 的概率选择“出 2 枚筹码”。此时，如果罗伊也按最佳策略行事，每局克林平均获利 1/12。假如罗伊不按最佳策略行事，罗伊的损失更大。这些结论都可以利用以前我们已经熟悉的方法得出。

108

下面接受这个问题的几何解法。这个解法非常直观，不需要复杂的计算。考虑图 6.2：

图 6.2



竖线 a 表示罗伊始终选择“出 1 枚筹码”，竖线 b 表示罗伊始终选择“出 2 枚筹码”。斜线 A 表示克林始终选择“出 1 枚筹码”，斜线 B 表示克林始终选择“出 2 枚筹码”，斜线 C 表示克林始终选择“出 3 枚筹码”。A 和 a 的交点表示当罗伊选择“出 1 枚筹码”而克林选择“出 1 枚筹码”时罗伊的收入，对应的值是 2；A 和 b 的交点表示当罗伊选择“出 2 枚筹码”而克林选择“出 1 枚筹码”时罗伊的收入，对应的值是 -3。B、C 与 a、b 的交点的含义类似。在 a 和 b 之间画一条竖线 c，c 表示罗伊以一定的概率选择“出 1 枚筹码”。c 与 A、B、C 相交，三个交点的含义是在罗伊的这个策略下，对应于克林的三种可能选择罗伊的收入。这三个交点决定了三个值，其中最小的一个代表罗伊在这个策略下的最少收入。c 从左向右移

动，由 a 的位置开始，到 b 的位置结束。在这个过程中，罗伊的最小收入不断变化，而最小收入有一个最大值。这个最大值对应的点称为“最佳点”。最佳点对应着罗伊的最佳策略，最佳点的高度表示罗伊在游戏中的平均收益，最佳点与 a 的距离比最佳点与 b 的距离的比例就是罗伊的最佳策略中两种选择的比例。

斜线 A、B、C 相交，有三个交点。最佳点必然是这三个点中的一个。比较这三个点即可找到最佳点。最佳点是两条斜线的交点，一旦发现最佳点，就可以把另外的一条斜线删去，这相当于删除克林的一种选择。此时问题转化为双方都只有两种选择的情况，依照以前的算法解决即可。

如果你用绘图纸精确地画出图 6.2，你甚至可以直接从图中测量出罗伊的最佳策略和平均收益。图中的最佳点已经用“*”表出，这一点的高度是 - 1/12，与 a、b 两条线的距离之比是 5:7。

除了直观以外，几何解法还有一个好处。如果游戏中存在109 优先选择，在图形上一目了然：图中存在两条不相交的斜线，其中一条始终在另一条之上。这个性质可以用来简化计算。

以上讨论的两种方法可以推广到克林的选择更多的情况。假设克林有 4 枚筹码，每次可以有 4 种选择，新增的选择是“出 4 枚筹码”。现在，克林的混合策略也多了三种，除了（1，2）、（1，3）和（2，3）以外，还有（1，4）、（2，4）和（4，5）。然而，问题并不因此而复杂。事实上，先前的解法依然有效，只不过计算的过程繁琐了一些。这个问题的细节留给读者完成。

现在克林多了一种选择：“出 4 枚筹码”，如果罗伊还是坚持以前的最佳策略，以 7/12 的概率出 1 枚筹码，那是不明智

的。因为假如罗伊以 7/12 的概率出 1 枚筹码，而克林出 4 枚筹码，克林的收益上升，平均每局高达 5/12。此时罗伊的最佳策略已经不同。罗伊应当以 9/16 的概率出 1 枚筹码，其平均收益是 - 3/16。如果克林在每局之前付给罗伊数额相当于 3/16 的钱，则游戏是公平的。

克林应当如何行事？他可以选择出 1 枚或 4 枚筹码，但不能选择出 2 枚或 3 枚筹码。一旦克林选择出 2 枚或 3 枚筹码，罗伊就可以获利。克林的最佳策略是以 11/16 的概率出 1 枚筹码，以 5/16 的概率出 4 枚筹码。

这个问题用几何方法解决也很方便。在图 6.2 中加入一条斜线 D，D 连接竖线 a 上的 - 5 和竖线 b 上的 + 6。D 与 A 的交点是最佳点，其高度为 - 3/16，与 a 和 b 的距离之比是 9:7。

在得出结论之后，你还需要一个产生随机数的器械（比如电子秒表）。如果没有这样的器械，即使知道了最佳策略你也无法应用。假如你没有辅助器械而对手有，对手就可以占便宜。 110

动物之间的游戏

大家都知道数学模型在物理学中的重要性，而数学模型对于生物学的重要意义则鲜为人知。其实许多重要的生物学理论背后都有美妙的数学理论的支持，比如解释生物遗传规律的孟德尔定律。现在流行的趋势是把对策论的原理和方法应用到对动物行为的研究中。

当动物之间争夺食物、领地、配偶等对于子孙繁衍至关重要的资源时，其行为可以当一个游戏来研究。我们考虑最简单

的情况，每个动物只有两种选择：鹰式策略和鸽式策略。所谓鹰式策略是指主动攻击，遇到攻击时坚决反击，除非遭受严重损失，否则绝不退让；所谓鸽式策略是指虚张声势，绝不展开真正的进攻，如果遭到攻击立刻撤退。两只鸽之间的决斗方式是张牙舞爪，徒劳地耗费时间和体能，直到有一方体力不支，退出竞争，另一方赢得胜利。

动物在竞争中获得的战果可以用"新增适应性"这个概念表达。所谓新增适应性即动物对环境的适应性增加，这个概念可以量化地描述，表示为因游戏的胜利而使自己繁衍后代的数量的平均值增加的数量。失败的一方也可以繁衍后代，但是后代的平均值会因失败而减少。新增适应性反映的不是后代的平均数量本身，而是整个平均数量的增加或减少。如果一个动物在游戏中选择的策略正确，其后代会增加，根据遗传的性质我们可以假设其后代继承它的策略。所以，如果鹰式策略优于鸽式策略，经过一段时间之后鹰式策略被选择的频率将上升。

为了给出支付矩阵，我们必须为"新增适应性"确立量化的标准和尺度。我们需要考虑的因素包括双方争夺的资源本身的价值，两个持鹰式策略的动物进行搏斗带来的损伤，两个持鸽式策略的动物耗费的时间和体能等等。以 10 表示资源的价值，假设两个动物都持鹰式策略或都持鸽式策略时双方获胜的机会均等，即平均每方可得 5，资源通过搏斗划分。如果一方持鹰式策略而另一方持鸽式策略，则后者望风而逃，前者得 10，后者得 0。此外，我们要考虑搏斗带来的损失。两个持鸽式策略的动物的搏斗耗费双方的时间和体能，每一方因此损失 2；两个持鹰式策略的动物的搏斗残酷而危险，这样的搏斗一旦发生，造成的损失已超过 5，此时双方各得 -2。支付矩阵如下：

	鹰式策略	鸽式策略
鹰式策略	-2	10
鸽式策略	0	3

现在我们要考虑的问题与先前完全不同。动物不会比较两个策略之间的优劣，所以我们关心的是**事实上哪个策略更好**，而不是如何选择更好的策略。我们必须采取新的研究方案。假设一个环境中的所有动物分为两类：鹰和鸽，前者所占的比例为 p，后者所占的比例为 1 - p。我们看一下哪一种动物对环境的适应性更好，适应性更好的一方后代会增加。

当一只鸽参与游戏时，其对手为鹰的概率是 p，此时其收益为 0；其对手为鸽的概率是 1 - p，此时其收益为 3。所以鸽的平均收益为

$$(0)\times(p)+(3)\times(1-p)=3-3p$$

利用同样的方法可以得出一只鹰参与游戏时的平均收益：

$$(-2)\times(p)+(10)\times(1-p)=10-12p$$

显然，当 $p=7/9$ 时，$3-3p=10-12p$。此时鹰式策略与鸽式策略一样好。当鹰所占的比例高于7/9时，$3-3p>10-12p$，即鹰式策略不如鸽式策略好（此时发生残酷搏斗的危险太高），鹰的后代数量会减少；相反，当鹰所占的比例低于7/9时，$3-3p<10-12p$，即鹰式策略比鸽式策略好，鹰的后代数量会增加。

经过很长很长时间之后，自然选择的结果会使一个动物采取鹰式策略的频率稳定在 7/9 上。一种极端情况是所有动物都只选择一种策略，其中的 7/9 每次都持鹰式策略，另外 2/9 每

次都持鸽式策略。另一种极端情况是每只动物都选择混合策略，以 7/9 的概率选择鹰式策略，以 2/9 的概率选择鸽式策略。无论如何，当一只动物参与游戏时，它的对手持鹰式策略的概率是 7/9，持鸽式策略的概率是 2/9。此时，无论这只动物持哪一种策略，其平均收益是一样的。现在考虑这个问题：如果在这样一个稳定的环境中，出现了一个突变物种，此物种选择的策略与众不同。此物种是否可以生存？

假设突变物种以 x 的概率选择鹰式策略，以 1 - x 的概率选择鸽式策略，x 是一个在 0 和 1 之间的数，不等于 7/9。当突变物种在环境中所占的比例很低时，它们在绝大多数的情况下与正常物种竞争。由于对手总是以（7/9，2/9）的比例选择 112 策略，所以无论突变物种选择哪种策略，其平均收益是一样的。突变物种的策略既不使其获得额外的好处，也不使其遭受额外的损失。这个物种生存还是灭亡，命运由环境随机决定。

但是当突变物种在环境中所占的比例增加到一定程度时，竞争发生在两只同属于突变物种的动物之间的概率上升。这样的两只动物在竞争中的平均收益决定了物种的命运。由于 x 不等于 7/9，动物在竞争中的平均收益下降。这说明，随着突变物种数量的增加，其对环境的适应性下降。无论此物种的成员以多大的概率选择鹰式策略，这个物种最终将被消灭。

J·M·斯密斯发明了一个专门的术语：进化稳定策略。简要地说，进化稳定策略的功能是抵制加入环境中的新物种。J·M·斯密斯在《进化与对策论》一书中介绍了许多类似的例子。在我们的例子中，以 7:2 的比例应用这两种策略就是进化稳定策略。

改变支付矩阵中的数值，进化稳定策略也随之改变。假设资源的价值下降而鹰和鹰之间搏斗的危险程度上升，支付矩阵

变成这样：

	鹰式策略	鸽式策略
鹰式策略	-4	6
鸽式策略	0	3

计算方法和以前一样。鹰和鸽在竞争中的平均收益分别是 6 - 10p 和 2 - 2p，当 p = 1/2 时，这两个值相等。在这个环境中，如果出现一个突变物种，其成员选择两种策略的比例不是 1:1，这个物种将被淘汰。以 7:2 的比例应用两种策略是这个环境中的进化稳定策略。

换一个例子。假设鹰和鹰之间的搏斗风险降低，支付矩阵变成这样：

	鹰式策略	鸽式策略
鹰式策略	1	6
鸽式策略	0	3

113

此时鹰和鸽在竞争中的平均收益分别是 6 - 5p 和 2 - 2p。由于 p 在 0 和 1 之间，所以 6 - 5p 总是大于 2 - 2p。这说明鹰式策略总是优于鸽式策略，鸽将被消灭，环境中的所有成员都是鹰。一个持不同策略的新物种无法加入这个环境。鹰式策略是进化稳定策略。

相反，假设鹰和鹰之间的搏斗极其残酷，一场搏斗使双方各得 -100。此时鹰所占的比例将剧烈下降，但不会灭绝。鹰

可以在一个很小的比例下生存，因为一只鹰与另一只鹰相遇的概率很低，而当鹰与鸽相遇时鹰还是占上风。在这个环境中，鹰和鸽的比例是 1:25。如果一个新物种选择两种策略的比例不是 1:25，新物种在环境中无法生存。1:25 是进化稳定策略。

当每只动物只有两种可选择的策略时，发现进化稳定策略是否由混合策略构成非常容易。观察一个支付矩阵，如果左下角的数值比左上角的数值大，并且右上角的数值比右下角的数值大，则进化稳定策略由混合策略构成；否则进化稳定策略由单一策略构成。如果进化稳定策略由混合策略构成，计算这个混合策略非常简单，和前文讨论的处理表 6.2 的方法很相似。以本节的第一个支付矩阵为例：

	鹰式策略	鸽式策略
鹰式策略	−2	10
鸽式策略	0	3

用第一行数值减去第二行数值，得到（−2，7）；删去正负号，交换位置，得到（7，2）；7:2 就是进化稳定策略。在第二个支付矩阵中，第一行数值减去第二行数值，得到（−4，4）；删去正负号，交换位置，得到（4，4）；1:1 就是进化稳定策略。在第三个支付矩阵中，第一行数值减去第二行数值，得到（1，4），这两个数都是正数，所以进化稳定策略不是混合策略。

114 动物不知道如何计算，但是自然选择的结果使得动物的行为与我们的计算结果相符。最关键的因素在于选择进化稳定策略可以使动物对环境的适应性增加，田纳森称之为“爪和牙的

先天红色”。鹰式策略和鸽式策略可以混合使用，利用数学方法可以计算环境的变化（资源价值的变化、动物行为方式的改变等等）如何影响两种策略的使用频率。如果鹰能想办法使两只鹰之间的搏斗造成的损失下降，对鹰的生存有利。

其实进化的过程已经实现了这个要求。在很多情况下，两只动物在开始搏斗之前先用某种方法比较双方的实力，比如显示自己的肌肉，或通过咆哮展示自己的肺活量。这与鸽和鸽之间的虚张声势不同，因为双方确实准备在必要的情况下战斗。如果某一方明显处于劣势，则战斗可以避免。这样对双方都有好处。T·C·布洛克介绍过一个精妙的例子。布洛克和同事们观察红鹿的行为。在交配季节，一只雄鹿要控制一群雌鹿作为自己的私有财产，这样有利于自己的基因繁衍。通常不到 7 岁的雄鹿身体不够强壮，没有能力控制雌鹿。而超过 11 岁以后雄鹿的战斗力下降。7 到 11 岁的雄鹿拥有行为能力，交配的欲望也很强烈。当两只雄鹿因争夺配偶发生冲突时，并不立刻开始战斗。通常，刚开始时两只雄鹿相距 100 米左右，向对方咆哮。它们可能是通过这种方式比较哪一方更强壮。动物不会使用扬声器，嘹亮的吼声无法伪造。如果双方在这个项目上不能决出胜负，双方会靠近，相距几米，做出侧移、跳跃等一系列动作，凶狠地晃动鹿角。这种行为似乎是向对方发出恐吓。观察者发现，这种预热活动进行的时间越长，战斗爆发的概率就越大。如果双方实力相差悬殊，通过吼叫、跳跃等行为可以鉴别出来。只有在预赛阶段无法分出胜负的情况下，残酷的搏斗才开始进行。

本章的主题是“抉择较少的游戏”。当每一方的选择只有两种时，进化稳定策略总是存在的。进化稳定策略有时候由混合策略构成，有时候则由单一策略构成。当每一方的选择不只

两种时，可能有很多种结果。有时候进化稳定策略包含各种选择，有时候进化稳定策略根本不存在。

两性博弈

雄性动物和雌性动物可以用相同的标准评价自己在繁衍后代方面的成就：看看自己有多少后代发育成熟，具备了传播祖先的基因的能力。生育出来的孩子如果在发育成熟之前夭折，则这次生育活动对于基因而言是没有意义的。理查德·道庭提出了一个简化模型，研究雄性动物和雌性动物在生育活动中的行为。在这个模型中，雌性可以选择两种策略：谨慎型策略和轻浮型策略，雄性也可以选择两种策略：责任型策略和放荡型策略。

谨慎型雌性的特征是只有当雄性愿意承担抚养后代的义务并取得自己的信任之后，才同意交配。谨慎型雌性的目的不是生育尽量多的后代，而是提高生育质量，使后代长大、成熟的条件更好，因为后代可以在双亲的监护下成长。谨慎型策略的缺点是如果对方不是责任型雄性，则交配无法完成，她会失去很多交配机会。

放荡型雌性的特征是不放过一切交配的机会，尽量多地繁殖后代。她不期望对方承担抚养后代的义务，但是如果真的遇到一个责任型的雄性，对她来说是一个额外的收益。

在这个游戏中，支付矩阵中的数值以成功地抚育出的成熟后代为标准。具体规定如下：每生育一次后代，父母各得 15；抚养一个后代长大需要耗费 20；如果父亲是责任型的，抚养责任由双方分担，即父母各付出 10。当谨慎型雌性与责任型

雄性相遇时，双方要耗费一定时间建立信任关系，这个“谈恋爱”的过程使双方各付出 3。当谨慎型雌性与放荡型雄性相遇时，交配不会发生，双方没有任何损失，也没有任何收益。雌性的支付矩阵如下：

	责任型雄性	放荡型雄性
谨慎型雌性	2	0
轻浮型雌性	5	-5

假设雄性中责任型所占的比例是 x。此时，谨慎型雌性的平均收益是 $2x$，而轻浮型雌性的平均收益是 $5x-5(1-x)=10x-5$。当 $x=5/8$ 时，这两个值相等。如果责任型雄性所占的比例超过 5/8，则轻浮型雌性有利；如果责任型雄性所占的比例低于 5/8，则谨慎型雌性有利。

从雄性的角度看，责任型的缺点是需要承担抚育后代的义务，而放荡型的缺点是可能遭遇最坏的情况：如果与谨慎型雌性相遇交配无法完成。雄性的支付矩阵如下：

116

	谨慎型雌性	轻浮型雌性
责任型雄性	2	5
放荡型雄性	0	15

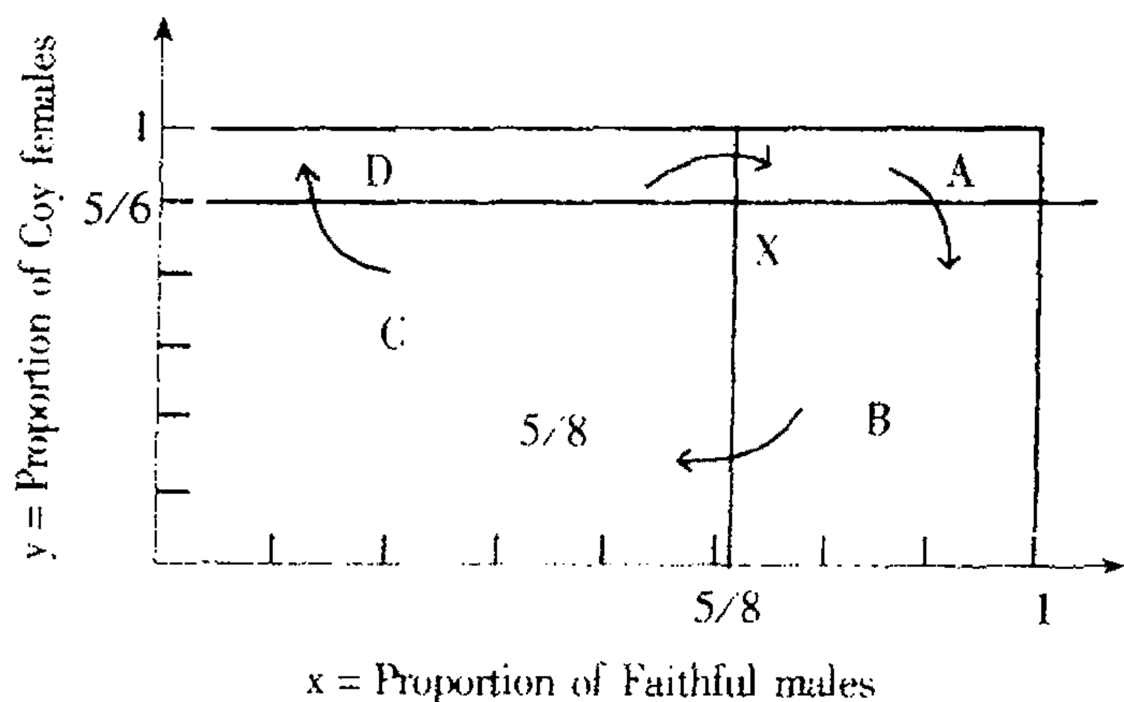
假设雌性中谨慎型所占的比例是 y。此时，责任型雄性的平均收益是 $2y+5(1-y)=5-3y$，而放荡型雄性的平均收益是 $15-15y$。当 $y=5/6$ 时，这两个值相等。如果谨慎型雌性所占

占的比例超过 5/6，则责任型雄性有利；如果谨慎型雌性所占的比例低于 5/6，则放荡型雄性有利。

x 的值和 y 的值相互干扰，并构成一个循环。如图 6.3 所示，图中的横坐标代表责任型雄性所占的比例，纵坐标代表谨慎型雌性所占的比例，图中任意一点代表一种全体成员的构成比例。如果 5/8 的雄性是责任型的，而且 5/6 的雌性是谨慎型的，则成员构成比例对应于 X 点，此时构成比例保持稳定。如果责任型雄性和谨慎型雌性所占的比例上升，成员构成比例对应于图中 A 区的某一点。在 A 区，由于责任型雄性的比例很高，轻浮型雌性占便宜，结果轻浮型雌性的后代增加，谨慎型雌性所占的比例下降，成员构成比例因而改变，对应于图中 B 区的某一点。

在 B 区，谨慎型雌性比例较低，而责任型雄性比例较高，形势对放荡型雄性有利，因为放荡型雄性很容易找到轻浮型的对象繁衍后代。结果放荡型雄性的基因传播迅速，责任型雄性所占的比例下降，成员构成比例因而改变，对应于 C 区的某一点。在 C 区，放荡型雄性比例较高，形势对轻浮型雌性不利，因为轻浮型雌性需要独力抚养后代。相反，谨慎型雌性总是在雄性的帮助下生育，形势对谨慎型雌性有利。结果谨慎型雌性的比例增加，成员构成比例因而改变，对应于 D 区的某一点。在 D 区，轻浮型雌性比例较低，放荡型雄性不易找到交配机会，形势对责任型雄性有利。结果放荡型雄性所占的比例下降，成员构成比例因而改变，重新回到 A 区的某一点。此后，下一轮循环开始。

在这个模型中，没有固定的最佳策略。你应当如何选择取决于异性的行为特征。雌性总是喜欢责任型雄性，雄性总是喜欢轻浮型雌性。雌性的品位决定雄性的行为。人类和动物有什



117

么区别?!

附 言

表 6.5 介绍了表 6.2 所代表的游戏进行若干局之后罗伊盈利的概率，其根据是计算机模拟游戏过程的结果而非严格的数学推理。下面我们讨论其中的某些细节，并与民意测验相比较。我的计算机内置了一个随机数发生器，根据它生成的随机数可以决定在一局具体的游戏中双方的选择（正面或反面），从而确定在这一局中每一方的收支情况。我们可以让计算机模拟若干局游戏，比如说 10 局，以这 10 局游戏的结果作为一组统计样本。在一组样本中，罗伊的最好结果是赢得 20 英镑，最坏结果是输掉 30 英镑。由于我们的目标是确定罗伊在若干局比赛之后盈利的概率，而不关心具体的盈利额，所以只需要判断在一组统计样本中罗伊的总收益额为正数、0 还是负数。不过，统计具体的盈利额也非全无用处，因为这个值可以用来检验我们的模拟程序是否工作正常。根据理论计算，罗伊在每

局中的平均收益为 1/8 英镑，在一组统计样本（10 局）中的平均收益为 10/8 = 1.25 英镑。我们可以生成许多组统计样本，计算出在每一组样本中的平均收益，看看结果是否接近于 1.25。当然，一组统计样本所包含的局数不一定是 10 局，这个数可以任意规定，50 局、200 局等等，均无不可。

具体规定多少局为一组统计样本取决于我们对统计精度的要求。在民意测验中，通常以 1 500 个调查对象为一组统计样本。这个数据广度有可能产生“样本偏差”。所谓样本偏差是指由于样本中所包含的调查对象不能有效地代表全部的调查对象，从而造成的误差。（除样本误差以外，还有许多因素可能导致结论失真。例如，调查对象对调查漠不关心或故意说谎，调查对象具备某些调查者未意识到的共同特征——都是未婚女性，或都已退休等等。在计算机模拟中不存在这些问题。）在实际操作中，我们不可能完全避免样本偏差，但通过增加数据广度，可以减少样本偏差。

为完成一次抽样调查总要有一些花费，直接消耗由两部分费用构成：调查的启动费用以及在调查过程中的费用。在民意测验中，我们直接考虑的费用是训练调查员的花费和付给调查员的工资。在计算机模拟中，通常不需要考虑这些问题。计算机是廉价而高效的工作人员，调查费用主要是编写计算机程序的花费，在程序给出以后，计算机完成 1 000 次调查的费用与完成 100 万次调查的费用差不多。计算机模拟通常可以产生极高的统计精度。在处理概率问题时，我们经常通过计算机模拟检验计算结果。如果我们的计算与实际结果有较大出入，通过计算机模拟很容易发现。

在民意测验中，即使我们能够保证样本的选择完全均匀，样本偏差依然无法避免。假设某个政党的实际支持率为 40%，

当我们以 1 500 人为抽样对象进行民意测验时，误差不超过一个百分点的概率为 2/3，而误差超过 2.5 个百分点的概率为 1/20。如果我们高估了某一个政党的支持率，同时我们也就低估了其反对党的支持率，所以当我们考虑两个政党的支持率的差的时候，误差还要大一倍。

民意测验的组织者可以通过增加调查对象的数目减少误差。然而，增加精度不是一件容易的事：为了使精度增加一倍，调查对象要达到原来的四倍。在计算机模拟中，我们只需要支付一些电费，可是在民意测验中，花费就大多了。

小　结

许多问题可以作为“抉择较少的游戏”来讨论。其中某些直接来自于游戏，比如猜硬币的游戏和出筹码的游戏，支付矩阵很容易获得。另外一些问题来自于社会冲突，为了得到支付矩阵需要慎重的考虑，而且对阵的双方对于支付矩阵应当是什么样观点可能完全不同。不过你可以不考虑观点上的分歧，仅按你的想法生成支付矩阵。为了得到你的最佳策略，关键在于寻找“最大极小值”。在所有游戏中，我们假设双方不知道对手将如何决策。 119

解决问题的第一步是检查“优先选择”是否存在。如果存在，则删除不必要的选择，这样可以使问题简化。下一步是检查鞍点是否存在。如果存在，则问题已经解决，游戏的结局是可以预测的。

当双方都至少有两个选择，而且鞍点不存在时，双方的最佳策略都由混合策略构成。在每一方的选择都多于两种时，问

题的解决非常复杂，我在本书中没有介绍。我们只研究至少一方的选择只有两种的情况。

最简单的情况是游戏双方都只有两种选择。第一步：用第一列（或行）数减去第二列（或行）数；第二步：删去正负号，交换位置；第三步：两个数的比例就是最佳策略中的两种选择的比例。利用这个结果可以求出游戏的公平参数，即罗伊的平均收益。如果每一局开始之前罗伊付给对手相当于公平参数的钱，游戏就是公平的。

当罗伊只有两种选择而克林的选择不只两种时，问题可以还原为双方都只有两种选择的情况。最关键的原则是克林的最佳策略一定是由不超过两种选择构成的。分别考虑克林可能采取的两种选择的组合，计算出罗伊在此形势下的最佳策略，以及这种策略对应的罗伊的最少收益。比较各个最小收益的值，其中最大的一个就是罗伊在这个游戏中的平均收益。罗伊的平均收益对应的双方的最佳策略就是在这个游戏中双方的最佳策略。这类问题也可以用几何方法解决。

如果问题反过来，罗伊的选择多于两种，而克林只有两种选择，解法是一样的。只要交换罗伊和克林在支付矩阵中的位置就可以了。不过不要忘记，原来的支付矩阵是从罗伊的立场制定的，所以支付矩阵中的数据应当调整，正数应当变成负数，负数应当变成正数。如果你的对手不知道（或不相信）我
120 们求最佳策略的方法，你需要做的仅仅是在实战中教训他。

在双方都只有两种选择的游戏中，如果一方始终选择最佳策略，而另一方不选择最佳策略，结果如何？非常有趣的情况，结果和双方都采取最佳策略的情况相同。比如在表 6.2 表示的游戏中，罗伊坚持以 3/8 的概率选择正面，则不论克林如何选择结果都是一样的。然而，如果克林偏离了最佳策略而罗

伊发现了这种情况，罗伊就可以调整自己的策略，从而获得好处。（比如克林每次都选择正面，则罗伊使自己选择正面的频率增加就可以获利。）

假设罗伊的选择只有两种而克林的选择有三种或更多。罗伊的最佳策略是混合策略，克林的最佳策略也是由两种选择构成的混合策略。如果罗伊采取最佳策略，而克林随机地选择两种策略中的一种，只要这两种策略是构成克林的最佳策略的成员，不论克林以什么样的概率选择，结果都是一样的。但是，一旦克林采取了第三种选择，罗伊就可以占便宜。罗伊不必为克林的愚蠢负责。

这些理论可以用来研究动物的行为。动物之间争夺那些与种族繁衍息息相关的资源。如果动物可选择的策略只有两种，则两种策略可以共存，并按一定比例形成进化稳定策略。进化稳定策略可以抑制新策略的产生。这种策略不关心资源的分配是否对种族的繁衍合理，它的功能只是消灭突变物种。当动物可选择的策略不只两种时，进化稳定策略可能不存在。

练习题

1. 支付矩阵如下，找出其中的最大极小值和最小极大值。鞍点是否存在？如果把表中的“15”换成“6”，鞍点是否依然存在？

	A	B	C
a	12	4	15
b	8	9	7
c	4	10	3

121

2. 计算表 6.6 中罗伊和克林的最佳策略。

3. 在习题 1 中，是否有优先选择？把表中的“10”换成什么数可以使 c 变成可删除的选择？

4. 如果删除习题 1 中的选择 A 和 c，双方的最佳策略是什么？游戏的公平参数是多少？

5. 设计一个方案，利用一枚骰子产生随机数。要求产生的随机数在 1 到 10 之间，每个数对应的概率是 1/10。

6. 游戏规则如下：罗伊和克林各有 4 枚硬币，每次双方可以选择出 0、1、2、3 或 4 枚硬币。胜负由双方出的硬币的总数决定。如果总数为 0、1、2、6、7 或 8，罗伊获胜；如果总数为 3，4 或 5，克林获胜。负者付给胜者 1 英镑。

画出罗伊的支付矩阵。删除其中的不必要选择。双方的最佳策略是什么？游戏是否公平？

7. 修改习题 6 的规则，规定如果总数为 0、1、7 或 8，罗伊获胜；如果总数为 2、3、4、5 或 6，克林获胜。结果有何变化？

8. 在鹰与鸽的对策模型中，假设资源的总价值是 12，两只鹰之间的搏斗使双方各损失 8，两只鸽之间的争斗使双方各损失 3。画出支付矩阵，求进化稳定策略。

9. 如果习题 8 中资源的总价值大于 12，结果如何？如果小于 12，结果如何？检验你的结论是否合理。

122

第七章　等待，等待，再等待

本章研究的主题是某件事有一系列的机会发生的情况。在某些问题中，我们关心预期的事需要等多长时间才能发生；在另一些问题中，我们关心预期的事发生多少次。在每个问题提出之后，我建议你首先独立地考虑一下，得出自己的结论或猜测，然后再继续阅读，检验自己的想法是否正确。读这本书需要主动地思考，如果你像读小说那样读这本书，恐怕不会有什么收获。

生日问题

有一群人，总数为N，为了使其中至少有两个人生日相同的概率大于二分之一，N至少是多少？这个问题与这群人是如何选择的有关。假如有一群双生子在接受心理试验，你闯进去把这群人作为调查对象，你会发现有两个以上的人生日相同的概率极高。在本题中，我们假设这群人是随机选择的，每个人的生日是一年中的任何一天的概率相同。

有很多因数会影响我们的结论。夏天出生的人比冬天出生的人多，润年的 2 月有 29 天。如果把这些因素都考虑进去，问题会变得非常复杂。数学家常用的策略是尽量简化问题。我们首先考虑最简单的情况：假设一年只有 365 天，不考虑润年的存在，而且每天产生的人一样多。在我们解决这个比较简单的问题以后，再把其它的因素考虑进去。

123

我们先计算任何两个人的生日都不同的概率。算出这个概率之后，只要用 1 减去这个概率就得到至少有两个人生日相同的概率。

当 N = 2 时，这群人只包括两个人。第二个人的生日是 365 天中的一天，为了使第二个人的生日与第一个人不同，他的生日可以是 364 天中的任何一天。因此，这个两个人的生日不同的概率是 364/365。

当 N = 3 时，前两个人的生日不同的概率是 364/365。为了使第三个人的生日与前两个人不同，他的生日可以是 363 天中的一天。因此，这三个人的生日各不相同的概率是

$$(364/365) \times (363/365)$$

类似地，当 N = 4 时，这四个人的生日各不相同的概率是

$$(364/365) \times (363/365) \times (362/365)$$

当 N = 5 时，这五个人的生日各不相同的概率是

$$(364/365) \times (363/365) \times (362/365) \times (361/365)$$

为了得到 N = 6、7 时的结果，只需把这个乘以 360/365，359/365，以此类推。

为了便于表达，我们用一个简单的表达式 $(364)_4$ 来表示 $364 \times 363 \times 362 \times 361$。这样，N = 5 时的结果就可以表示为

$$(364)_4/365^4$$

对于任意的 N，这 N 个人的生日各不相同的概率是

$$(364)_N/365^N$$

当 N 比较小时，这个值接近于 1。随着 N 的增加，这个值递增。当 N 达到 366 时，这个值等于 0。于是，我们最初的问题变为“当 N 达到多大时，$(364)_N/365^N$ 小于 50%”。 124

你认为答案应该是多少？除非以前见过这个问题，多数人会把答案估计得很高。很多人选择 183，因为这个数刚好大于 365 的一半。很少有人选择 40 以下的数。正确的答案是 23。

当 N = 22 时，$(364)_N/365^N$ 约等于 0.524 3；当 N = 23 时，$(364)_N/365^N$ 约等于 0.492 7。所以答案是 23。如果这个结论让你吃惊，我可以提供另外一个分析思路。

任何两个人生日相同的概率是 1/365。在 23 个人中任意取出两个，一个有 $^{23}C_2 = 253$ 种取法。因此，23 个人中有两个人生日相同的概率大致在 253 × （1/365） = 253/365 附近。在我们的计算过程中，有些情况被重复计算，比如 A 和 B 的生日相同，B 和 C 的生日相同，则 B 和 C 的生日一定也相同，这种情况被计算的次数不止 1 次。所以，最后的结果要比 253/365 小一些。无论如何，我想向你说明的是：当人数达到 23 时，至少有两个人生日相同的概率已经相当大了。至少比一般人估计的大得多。

现在我们可以把一些复杂的因素考虑进来。事实上，每天出生的人数是不同的。这个事实会对我们的结论产生什么影响？一个很重要的原则是：**当每个人的生日在 365 天中的分布不均匀时，至少有两个人生日相同的概率上升。**要证明这个结论不像想象中的那么复杂。方框中的文字是证明过程（没兴趣的读者可以略过这一段）。 125

当每个人的生日在 365 天中均匀分布时，23 个人中至少有两个人生日相同的概率大于二分之一。如果每个人的生日在

365天中的分布不均匀，则这个概率更大，也许22个人中至少有两个人生日相同的概率就已经大于二分之一了——这取决于分布的均匀程度。

生日分布不均匀时的生日问题

假设每个人的生日在365天中的分布是不均匀的，即存在一些日子，它们是某个人的生日的概率不等于平均值。在其中选出两天X和Y，使得X是某个人的生日的概率x大于平均值，而Y是某个人的生日的概率y小于平均值。为了计算这群人中任何两个人的生日都不同的概率，需要列出一个极其复杂的表达式。（当N等于4时，这个表达式包括7亿多个符号！）我们不关心这个表达式的细节，仅用A表示它。

我们把X和Y是某个人的生日的概率改为 $(x+y)/2$，并保持其它日子对应的概率不变。此时表达式A变为另外一个表达式，用B表示这个表达式。在A和B之间，只有那些和X与Y有关的项是不同的，其它项完全一样。所以计算A和B的差额很容易。事实上 $B-A$ 等于 $(x-y)^2/4$ 乘以一个系数。我们可以肯定这个系数是正数，因为它表示某一个概率。

因此，$B-A$ 一定大于0。即当每个人的生日在365天中的分布不均匀时，随着均匀程度的上升（把x和y都替换成 $(x+y)/2$ 使均匀程度增加），这群人中任何两个人的生日都不同的概率下降。当每个人的生日是任何一天的概率都等于1/365时，均匀程度最大，而任何两个人的生日都不同的概率最小。也就是说，至少有两个人生日相同的概率达到最大。

126

如果把2月29日考虑进来结果如何？假设每个人的生日在366天中均匀分布，则当N为22和23时，任何两个人的生日都不同的概率分别是0.525 4和0.493 7。所以结论不变，当

N = 23 时，至少有两个人的生日相同的概率大于 1/2。事实上，由于一个人的生日是 2 月 29 日的概率很低，所以润年的存在使得生日的分布变得不均匀，至少有两个人的生日相同的概率还要大一些。

一个名叫罗勃特·马修的人曾实际检验过这个结论。他的方法是统计一场足球赛上的 23 个人的生日（22 名场上队员，1 名裁判）。他统计了 1997 年 4 月 19 日举行的 10 场甲级联赛，其中 6 组统计结果中存在生日相同的人，另外 4 组统计结果中没有生日相同的人。理论和观察符合得很好。

我们的结论也可以在学校里检验。一个班级的学生人数从 26 人到 32 人不等。当 N = 26 时，至少有两个人生日相同的概率为 60%；当 N = 32 时，至少有两个人生日相同的概率为 75%；把两个人数为 25 的班级合并统计可以得出 N = 50 的结果，此时至少有两个人生日相同的概率超过 97%。

《破坏者》问题二

在上一章，我们从《破坏者》一书中选了一个例子。这个例子也来自《破坏者》。两个人——弗莱特和华伦——驾车旅行。为了排遣旅途的枯燥，弗莱特提议赌钱。具体的赌法是计录从对面开过来的汽车牌照号码的最后两个数字。如果连续 20 辆汽车中至少有两辆牌照号码的最后两位数字相同，则弗莱特胜；如果各不相同，则华伦胜。赔率是 1 比 1。

127

华伦觉得这是一个非常不错的机会。最后两个数字在 00 到 99 之间，共有 100 种可能。两辆车的牌照号码的最后两位相同的概率仅为 1/100。即使考虑连续 20 辆汽车，这个概率也

不过是 20/100。因此，自己的胜率高达 80%。多美妙的事啊！但实战的结果让华伦困惑不已：几乎每一轮都是弗莱特获胜！到达终点之后，弗莱特揭示了谜底。

这个问题与生日问题一样，20 辆汽车牌照号码的最后两位各不相同的概率是

$$(99)_{19}/100^{19}$$

约等于 0.130 4。也就是说，华伦的胜率仅为 13%，对手的胜率高达 87%。即使对方提供 1 比 5 的赔率，游戏也对华伦不利，更何况赔率是 1 比 1！

当选择的汽车辆数是 20 时，游戏明显对华伦不利。当我们把 20 替换成其它数时，这个游戏是否可以变成一个公平游戏？当替换为 11 时，华伦的胜率是 0.565；当替换为 12 时，华伦的胜率是 0.503；当替换为 13 时，华伦的胜率是 0.443。所以，如果把 20 替换为 12，这个游戏基本上是公平的。你可以在旅途中和别人玩这个游戏。

通用的计算方法

当结果可能有 365 种时，对象的个数为 23 使得至少出现一对相同结果的概率接近于 1/2；当结果可能有 100 种时，需要的对象的个数为 12。是否存在一种通用的计算方法可以用来方便地求出结论？以 N 表示结果可能有多少种，以 x 表示需要的对象的个数，则 x 约等于 $\sqrt{1.4N}$。当 N = 365 时，x 约等于 $\sqrt{511}$，这个数在 22.6 附近，所以我们的结论是 23。当 N = 100 时，x 约等于 $\sqrt{140}$，这个数在 11.8 附近，所以我们的结论是 12。如果你不想知道这个公式是怎么来的，可以略过证明

128

过程。

在大英帝国彩票中，头奖的号码组合共有 1 400 万种可能，经过多少轮出现两个相同的号码组合的概率接近于二分之一？用上面的近似公式计算，结果是 4 400。按每周抽彩两次计算，4 400 轮需要 42 年。

通用公式的推导过程

假设结果可能有 N 种，则任意一对对象结果相同的概率是 1/N，而结果不同的概率是 1 - 1/N。在 K 个对象中，任意选出两个组成一对，可以有 $^{K}C_2 = K(K-1)/2$ 种选择方法，用 M 表示 K（K - 1）/2。（我们没有考虑对象被重复计算的情况，不过没关系，忽略这种情况只会引起很小的误差。毕竟我们的目标是找出近似公式，绝对的精确性是不必要的。）

这 K 个对象结果各不相同的概率是 $(1 - 1/N)^M$，其中 M = K（K - 1）/2。当 N 是一个比较大的数时，$(1 - 1/N)^M$ 约等于 $e^{-M/N}$。为了使 $e^{-M/N}$ 接近于 1/2，M 应当等于 Nln（2）。即 K（K - 1）约等于 2Nln（2）。当 K 比较大时，K（K - 1）约等于 K 的平方，所以 K 的平方约等于 2Nln（2），于是我们得到估计值 $\sqrt{1.4N}$。

129

也许你需要一些更详尽的资料，比如当弗莱特向华伦提供一定比例的赔率时，x 取什么值才能使游戏公平？表 7.1 提供了几组数据。你可以利用这个表指导你的实战。

表 7.1

赔率	计算 x 的公式	当 N = 365 时的 x 值	当 N = 100 时的 x 值
1:1	$\sqrt{1.4N}$	23	12
2:1	$\sqrt{2.2N}$	28 ~ 29	15
3:1	$\sqrt{2.8N}$	32	17
6:1	$\sqrt{3.9N}$	38	20
10:1	$\sqrt{4.8N}$	42	22

多个人生日相同的概率

前文我们介绍过，一个名叫罗勃特·马修的人曾实际检验过生日问题的结论。在他的十次检验中，有两次检验的结果出现了有两对人生日相同的情况。这个频率是否正常？分析这个问题不会太复杂，方法类似于我们在上一节中推导通用公式的过程。数学家已经准备好一种成熟的计算方法来处理这类问题，在数学上称为“泊松分布”。我们先来计算每次检验中平均应发现多少人生日相同，然后生日相同的人数恰好为 x（x = 1，2，3，…）的概率。

在 K 个对象中选择两个，有 $^{K}C_2 = K(K-1)/2$ 种选择方法。这一对对象的结果是具体的某一种的概率为 1/N。（在生日问题中，对象即为人，结果即为生日，一对对象的结果是具体的某一种即两个人的生日是同一天，K = 23，N = 365。）所以，结果相同的对象有 K（K - 1）/2N 对，用 μ 表示这个值。在生日问题中，N = 365，K = 23，所以 $\mu = 0.693$。根据附录 2

中介绍的计算方法，在 10 次检验中，应当有 5 次检验不出现生日相同的人，有 3.5 次检验恰好出现一对生日相同的人，有 1.5 次检验出现的生日相同的人多于两对。在罗勃特·马修的实际检验中，有 4 次检验不出现生日相同的人，有 4 次检验出现一对生日相同的人，有 2 次检验出现两对生日相同的人。理论与实际符合得非常好，罗勃特·马修的数据是正常的。

出现三个人生日是同一天的概率是多少？在多少人中出现这种情况的概率接近于二分之一？计算方法是类似的。在 K 个人中，存在 K（K－1）（K－2）$/6N^2$ 个组，每个组包含三个生日相同的人。当 K 约等于 1.6 $N^{2/3}$时，出现三个人生日相同的概率接近于 50%。由于 N＝365，所以 K 约等于 82。即在 82 个人中，存在三个人生日相同的概率约等于 50%。如果把这个公式应用到猜汽车牌照的游戏中，则 N＝100，K 约等于 35。即在 35 辆汽车中，存在三辆汽车牌照后两位数字相同的概率为 50%。

如果我们考虑四个人生日是同一天的概率，得到的公式是 K 约等于 2 $N^{3/4}$。N 取 365 时，K 约等于 167。即在 167 个人中，存在四个人生日相同的概率约等于 50%。

扑克牌问题一

取两副扑克牌，一副放在左边，一副放在右边，分别洗乱。翻开两叠牌的头一张，如果两张牌一样，则称为“相配”；如果不一样，则称为“不相配”。然后翻开两叠牌的第二张，同样，如果两张牌一样，则称为“相配”；如果不一样，则称为“不相配”。然后再翻开两叠牌的第三张，检查是否相配，

131

然后再翻开第四张，如此等等，直到翻开第 52 张。[1] 有可能在全部的 52 次检验中一次相配也不发生，也有可能发生多次相配。现在的问题是：平均应当发生多少次相配？请读者考虑这个问题，如果你不知道如何计算，可以凭直觉猜一个答案。

我曾让许多人回答这个问题，多数人认为平均应当出现 4 次或 5 次相配，少数人认为低于两次，极少数人认为高于 5 次。下面我们计算一下。现考虑头一张牌发生相配的概率。无论左边的一叠牌头一张是什么，右边的一叠牌的头一张与之相配的概率是 1/52；其次考虑第二张牌，无论头一张牌是否相配，也无论左边一叠牌的第二张是什么，右边的一叠牌的第二张与之相配的概率也是 1/52。（事实上，如果我们已经知道了两叠牌的头一张是什么，对于第二张是否相配的概率是有影响的，但为计算简便我们可以忽略这种影响，而且对计算结果的精度影响不大。）同理，右边一叠牌的任何一张与左边相应的牌相配的概率都是 1/52。结论非常明显：一共有 52 种相配的可能，每种可能发生的概率是 1/52，所以平均应当发生 $51\times 1/52=1$ 次相配。

如果一副扑克牌的张数不是 52，以上结论依然有效。例如，每副牌只留下黑桃，则每叠牌只剩下 13 张。此时，发生相配的平均次数应当是 1。换一个例子，取两堆小球，每堆 15 个，分别标以从 1 到 15 的号码。把两堆小球装入袋子，一袋放在左边，一袋放在右边。同时从两个袋子中取小球，每次从两个袋子中各取一个，比较两个小球上的号码是否相同，如果相同则称为一次“相配”。连续取 15 次。平均应当出现多少次相配？答案是 $15\times 1/15=1$。

---

① 西方人的习惯是把两张王牌去掉，一副只有 52 张。——译者注。

这个问题本来非常简单，但是许多人把它想复杂了。最关键的原则是"一个和的平均值等于各个平均值的和"（参阅附录 3），掌握了这个原则问题就迎刃而解。

如果你想知道具体发生多少次的相配的精确概率，计算方法非常麻烦，不过利用泊松分布可以得到近似结果。当扑克牌包含的扑克张数较多时（比如大于 8），近似值的精确度非常理想。在附录 2 中可以找到具体的计算方法，下表是计算结果。

132

发生相配的次数	0	1	2	3	4	5 或更多
对应的概率（%）	37	37	18	6	1.5	0.4

如果你准备和别人赌发生相配的次数，这个表可以帮助你确定你能忍受的赔率。

以上方法可以应用于许多不同的问题。下面是三个例子。

1. 你写了 N 封信，并为每封信写好了信封。你把装信的任务交给粗心大意的秘书，但是在装信的过程中，她随机地为每封信选择一个信封。问题是平均有多少封信被装入了正确的信封？你当然知道答案：1 封。（类似的问题是在剧院门口的衣帽间，服务员胡乱地把帽子交给顾客，平均有多少个顾客拿到自己的帽子？）

2. M 对夫妇结伴旅游，旅馆随机地为他们分配房间，每间房两个人，一男一女。平均有多少对夫妇住在一起？

3. 约翰整理自己的书架，把书按作者的名字的字母顺序排列，作者相同（或同名）的书按出版时间排序。过了一段时间，他重新排列图书，按书名的字母顺序排序。无论他有多少

本书，平均而言有1本书位置保持不变。

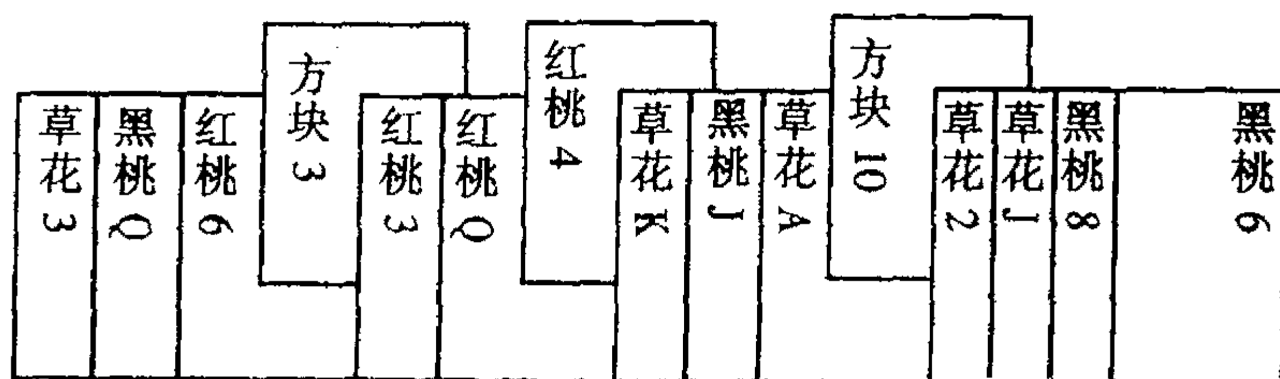
扑克牌问题二

这个问题的答案相当令人吃惊，不过假如你不小心，结果可能是灾难。你需要一群观众，在他们面前演示这个游戏。你在观众里选一个人，比方说莎拉。你取一副扑克牌，把顺序洗乱，然后平铺在桌面上。让莎拉在前10张牌中任意选一张，但不要告诉你她选的是哪一张。让莎拉在这张牌的背面做一个标记，然后从这张牌开始向后数牌，这张牌的号码是几就向后数几张，用这个办法挑出下一张牌；再在这张牌的背面做一个标记，然后再向后数牌，第二张做标记的牌号码是几就向后数几张，用这个办法挑出第三张做标记的牌。以此类推，直到挑出最后一张做标记的牌。10、J、Q、K的号码都当做10，A的号码当做1。如图7.1所示，假如莎拉选择的牌是第四张，方块3，则第一张做标记的牌就是方块3；向后数3张，得到红桃4，于是第二张做标记的牌就是红桃4；向后数4张，得到方块10，于是第三张做标记的牌就是方块10。按这种方法挑出最后一张做标记的牌。

显然，最后一张做标记的牌必须出现在最后10张牌中。假如第48张牌做了标记，而这张牌是7，那么这张牌就是最后一张做标记的牌；如果这张牌是2，则还可以挑出第50张牌，于是第48张就不是最后一张做标记的牌。你不许看莎拉挑牌的过程。你的任务是猜测最后一张做标记的牌是哪一张。你认为你猜中的概率是多少？

刚开始，你可能认为最后的10张牌中的每一张是最后一

图 7.1



张做标记的牌的概率是相等的，所以你猜中的概率是十分之一。不过如果你检查一下最后 10 张牌，你会发现有些牌不可能是最后一张做标记的牌。比如说，如果第 48 张牌是 2，这张牌就不可能是最后一张做标记的牌。类似地，第 45 张牌是 A，如果第 47 张牌是 4，这两张牌也不能作为候选答案。这样你可以使猜中的概率增加一点，也许可以达到六分之一或七分之一。可是你是否相信，我有办法使猜中的概率超过 70%！

怎样才能达到这么高的成功率呢？方法其实很简单，你需要做的就是在前十张牌中任意选一张，然后按照与莎拉同样的方法向后数，直到得到最后一张牌。只要以这张牌作为你的答案，你就可以保证极高的成功率。

为什么会这样呢？你事先不知道莎拉选哪张牌，所以你有 1/10 的机会与她的选择重合。此外，你得到的某张牌可能与她标记的某张牌重合，一旦这种情况发生，你们俩选出的最后一张牌肯定是一样的。你们俩在起点处不重合的概率是 9/10，但是你有很多次机会与她标记的某一张牌重合。只要重合发生一次，你就可以保证得到正确答案。

下面我们估算一下猜中的概率。莎拉从前 10 张牌中的某一张开始，所以平均而言她的起点是第五张牌。有五种牌会使莎拉向后数 10 张牌（10、J、Q 和 K），所以平均每次莎拉会向后数

6 到 7 张牌((1+2+3+…+9+10+10+10+10)/13)。一共有52张牌,所以平均而言莎拉选出8张牌(包括最初一张)。在你选择牌的时候,你每次也是跳过6到7张牌,你选择的结果与她选择的结果不重合的概率是6/7,连续8次都不重合的概率是$(6/7)^8$,约等于0.29。只要重合一次,你的最后结果就与她相同,所以你成功的概率高达71%。如果你还希望这个概率更大一些,你在选择起点的时候应当选择一张靠前的小牌,这样重合发生的概率进一步增加。举一个例子,如果开始的几张牌是7、10、4、Q、K、2……,你应当以4为起点。

以上分析相当粗糙,但是给出了一个很有道理的估计值。如果你想得到精确的计算,必须考虑全部52张牌的具体排列顺序,这几乎是不可能的。为了检验我们的结论是否可靠,应当拿一副扑克实际操作几次,不过最好的检验办法还是计算机模拟。计算机模拟的优点是可以非常迅速地产生大量的数据,而只有数据非常丰富时检验才有说服力。计算机自动生成一副牌,然后分别从前十张牌中的每一张开始向后标记牌,记录每一次的最终结果是哪一张牌。有时候这十次得到的结果是同一张牌,这意味着在这副牌中你可以保证成功率是100%。有时候这十次得到的结果不是同一张牌,不过计算成功率也是非常方便的。例如,十次中有七次终止于第46张牌,三次终止于第51张牌,则你的成功率是$0.7\times0.7+0.3\times0.3=58\%$。我让计算机模拟了1万副牌,其中有2 800副牌无论从哪张牌开始结果都一样。计算机统计出来的成功率略大于70%。这与我们的计算结果符合得非常理想。(在这个数据密度下,60~80%之间的任何结果都应被视为可接受的。)

135

你可以使成功率更高一点,只要你尽量选择靠前的小牌就可以了。莎拉通常会在第六张牌和第七张牌附近选择起点。如

果你们俩都在前5张牌中选择起点，则你的成功率可以高达75%。

当然，前提条件是你们两个都不能查错牌。如果某一人查错了牌，你的成功率将剧烈下降。所以你不但要保证自己不出差错，还要督促莎拉仔细从事。你可以先为大家演示一下标记牌的过程，然后让全体观众监督莎拉的操作。你的成功率肯定会让所有人震惊，不过偶尔你也会失手。为了使成功率更高，你可以把两副扑克牌混在一起做这个游戏。在扑克牌总数为104张的情况下，你的成功率超过90%。当然，在两副牌的情况下你犯错误的可能性也增加了，所以一定要小心从事。

搜集卡片

战前的香烟盒里通常有一张附赠的卡片，今天仍有人收集这些卡片，希望能得到完整的一套。卡片上通常是30年代著名球星的照片和简介。现在，许多小商品(如苞米花、糖果等)的包装盒里也有类似的卡片。如果你想搜集一套完整的卡片，平均你应购买多少件商品？ 136

和生日问题一样，我们首先要把问题限制得明确一些。为了便于讨论，我们假设厂家投放的各种卡片的数量一样多，而且随机地在每个包装盒里放入一张卡片。此外，我们不考虑与其他购买者交换卡片的可能。首先我们估计一下，如果厂家发行了24种卡片，平均我们需要买多少件商品才能收集到完整的一套卡片？另外，如果卡片的种类增加了一倍，我们需要购买的商品数是否也要增加一倍？问题不是一目了然的。

我们分步考虑你收集到新卡片的过程。当你买第二件商品

时,两件商品中的卡片相同的概率较低(仅为 1/24),所以你有 23/24 的机会得到第二张新卡片。刚开始你会经常得到新卡片,然而,随着你拥有的卡片的种类的增加,你购买一件新商品时得到新卡片的概率下降。也许你只买了 15 件商品就得到 12 种卡片,此时,为了得到第 13 种卡片你平均要买多少件商品?由于卡片一共有 24 种,你买下一件商品时,得到的卡片与旧卡片重复的概率是 12/24,所以你有 1/2 的概率得到新卡片。也就是说,为了得到第 13 种卡片,平均你要买 2 件商品。

如果你已经得到了 23 种卡片,为了得到最后一张卡片,你需要购买的商品将多得多。当你买下一件商品时,得到旧卡片的概率高达 23/24,得到新卡片的概率只有 1/24。所以,为了得到第 24 张卡片,平均你需要购买 24 件商品。

总之,刚开始时你很容易得到新卡片,而一旦你已经得到半数的卡片,收集过程就变得缓慢了。尤其在收集最后几张卡片时,往往需要漫长的等待。当你已经收集到 n 张卡片时($n=0,1,2,\cdots,23$),以 $x(n+1)$ 表示为了得到第 $n+1$ 张卡片所需要购买的商品数。显然,$X(1)=1$,$X(2)$略大于 1,$X(13)=2$,$X(24)=24$。你需要购买的商品总数为

$$X(1)+X(2)+X(3)+\cdots+X(24)$$

用 A 表示这个数。如果我们能找到一个通用的算法计算 $X(n)$,137 则问题得到解决。

我们首先考虑 $X(18)$。如果我们得到了 $X(18)$的计算方法,其他值的计算过程是类似的。此时,你已经得到了 17 种卡片,目标是得到第 18 种。由于卡片一共有 24 种,当你买下一件商品时,得到旧卡片的概率是 17/24,得到新卡片的概率是7/24。所以,为了得到第 18 种卡片,平均你需要购买 24/7 件商品。因此,$X(18)=24/7$。利用同样的方法可以求出各个值,$X(19)=$

24/6,X(20)=24/5,等等。所以

$$A=\frac{24}{24}+\frac{24}{23}+\frac{24}{22}+\cdots+\frac{24}{7}+\cdots+\frac{24}{24}+\frac{24}{2}+\frac{24}{1}$$

也就是说，$A=24\left(1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\cdots+\frac{1}{24}\right)$。这个值约等于 91。所以答案就是 91。

显然，当卡片一共有 48 种时以上解法依然有效，所以，为了得到全部的 48 张卡片，需要购买的商品总数为

$$A=48\left(1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\cdots+\frac{1}{48}\right)\approx 214$$

这个值比 91 的二倍大一些。

实际上我们已经得到了这一类问题的通用解法。当卡片一共有 N 种时，为了得到全部的 N 种卡片，平均需要购买的商品数为

$$P=N\left(1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\cdots+\frac{1}{N}\right)$$

在数学上，$1+1/2+1/3+1/4+\cdots+1/N$ 是一个很常见的公式，当 N 比较大时，计算这个值非常麻烦。但是不要紧，数学家早就准备好了方便的近似公式。当 N 比较大时，P 近似等于

$$N\{\ln(N)+0.577\}$$

138

这个公式的误差很小。当 N 大于 10 时，误差低于 2%。

我们可以在大英帝国彩票中应用这个公式。我们的问题是：为了使从 1 到 49 的每一个号码都曾经被作为额外机会号码，平均需要开奖多少轮？在这个问题中，N=49，平均需要开奖 219 轮。大英帝国彩票的实际情况是，最初的 30 轮产生了 26 个不同的额外机会号码，但此后产生新号码的过程缓慢得多。此后的 20 轮中只产生了 10 个新号码。到第 77 轮时，只有 5 个号码尚未作为额外机会号码出现过。理论计算的结果是，在第

77 轮以后产生下一个新号码需要 10 轮(49/5),实际上用了 36 轮。第 46 个号码的产生用了 16 轮,第 47 个号码的产生用了 26 轮。到第 155 轮时,只有两个号码尚未出现过。理论计算的结果是第 48 个号码需要 24.5 轮,而最后一个号码需要 49 轮。实际情况与计算略有出入:在此后的 100 轮中,第 48 个号码都没有出现。突然间,最后两个号码在连续两轮中出现了。整个过程用了 262 轮。这与理论计算的结果(219 轮)有出入,不过可以认为符合得比较好。

这个公式也可以用来解答这个问题:在大英帝国彩票中,为了使每一种号码组合都有人购买,平均需要卖出多少张彩票?在第二章中我说过,各种组合被购买的频率不同,但是在生日问题中我们已经得出结论:频率的不均匀分布将使得重复的结果增加。所以,实际的数字应该比我们的计算结果更大。在这个问题中,可选择的号码组合有 1 400 万种,即 N 等于 1 400 万。此时算出的 P 值为 2.4 亿。事实上,任何一轮彩票的销售量都不曾达到这个数的一半。这说明,在正常情况下,每一轮彩票中139都应当有未被购买的号码组合,除非有某个大财阀有意买进许多不同的号码组合。

为了使每一种号码组合都有人购买,为什么需要卖出如此多的彩票?你可以非常直观地考虑这个问题:假设所有号码组合中只剩下一个无人购买,还需要卖出多少张彩票才能把这个最后的号码组合卖出?当卖出下一张彩票时,其号码组合不与已卖出的号码组合重复的概率仅为 1 400 万分之 1,也就是说,为了卖出最后一个号码组合就需要再卖 1 400 万张。为了卖出最后几个号码组合,需要卖许多许多张彩票。

换一个例子。你站在马路边,问每一个路过的人他(或她)出生在哪个月。平均需要问多少人才能使 12 个月在答案中都

出现过？假设人的生日在12个月中均匀分布。由于N=12,结论是37。不过事实上人的生日在12个月中的分布不均匀(有的月份只有28天,有的月份有31天),所以你得到的重复答案比理论估计要多。所以正确的结论应比37大一些。我们最好把结论估计为40。

我们已经知道如何计算“为了得到全部N种卡片,平均需要购买的商品数”。另外一个问题也需要讨论:对于一个给定的x(x不小于N),买x件商品恰好收集到全部N种卡片的概率是多少？分析这个问题要用到附录3中介绍的几个概念,比如方差、标准差、正态分布等等。我们略过具体的解法,直接给出结论:

○ x大于平均值的概率与x小于平均值的概率相等;

○ x与平均值的差的绝对值不超过标准差的概率大约为2/3;

○ x与平均值的差的绝对值超过标准差的二倍的概率大约为1/20。

例如,当平均值为100而标准差为10时,x落在90到110之间的概率是2/3;x大于110的概率和x小于90的概率相等,都是1/6;x大于120或小于80的概率是1/20。如果平均值仍为100而标准差为30,则结论颇为不同:x大于130或者小于70的概率是1/3。

表 7.2

N	12	24	48
平均值	37	91	214
标准差	14	29	60
2/3 概率对应的区域	23 ~ 51	62 ~ 120	154 ~ 274

如果你想知道理论的细节，需要读专门的概率论教科书。附录 3 中有简介。为了应用以上结论，我们还应该知道如何计算标准差。当每一种卡片出现的频率相同，并且卡片的种类多于 10 种时，标准差的近似值为 $N\pi/\sqrt{6}$。(你可能会奇怪这个公式怎么和 π 扯上关系的。) 这个值可以很好地近似为 1.28N。表 7.2 给出了几组数据。

假如你真的开始收集卡片，并且发现实际结果和理论计算不一致，不值得大惊小怪。实际结果可能比平均值高很多，也可能低很多，这很正常。

以上分析基于一个共同的假设：每种卡片出现的频率相同。如果每张卡片出现的频率不同，结论会有变化：我们需要买更多的商品才能达到目标。此时，决定性的因素是我们平均需要买多少件商品才能得到最罕见的那张卡片。比如，一共有 12 种卡片，其中 11 种卡片出现的概率相同，而最后 1 种卡片出现的概率是 1/100。平均而言，我们买 100 件商品可以得到 1 张最稀有的卡片，而此时我们通常已经收集到全部的 12 种卡片——如果此时还有一些普通卡片我们没有得到，只能说明我们的运气非常非常差。刚才我们已计算过，平均用 37 轮就可以得到 12 种分布均匀的卡片，而在 65 轮之后依然没有达到目标的概率仅为 1/40。

你是不是觉得有点奇怪？频率分布不均匀的情况要比频率

分布均匀的情况复杂，可是计算复杂情况下的结果竟然比计算简单情况下的结果简单得多！如果在全部卡片中，只有一张卡片非常少见，其它卡片出现的频率比它高得多，则问题变成一个非常简单的形式：平均需要购买多少件商品才能得到一张最罕见的卡片？算出这个值以后，我们最后的结论要比这个值略大一点点。

141

选秘书问题

假设你要招聘一名秘书。有一群应聘者在你的办公室外排队等候面试。你依次面试这些人。每次面试结束时你有两个选择：1. 回绝这个应聘者，面试下一个人；2. 决定聘用当前的应聘者，结束招聘活动。你必须当场做出决定。被你回绝的人，你不能聘用。怎样才能使你聘用到素质最好的应聘者的概率最大？

其实生活中的某些例子与此相似。比如你来到一个陌生的小镇，要在一排小餐馆中选择一家吃午饭，如何选择最合理？在一间餐馆停下来就餐，还是继续往前走？

你的目标是使你聘用到素质最好的应聘者的概率最大。假设应聘者共60人，如果你胡乱选择一个，你有1/60的机会选对人。当然你有办法使机会更大。

通常的策略是这样：回绝最初的一部分应聘者，然后考察后来的应聘者，如果出现一个应聘者，其素质比以前的所有应聘者都好，则聘用此应聘者，招聘活动终止。最糟糕的情况是，你发现后来的应聘者的素质总是比最初的某个应聘者差，结果你不得不聘用最后一个应聘者——可能是很差劲的一个。

现在，最关键的问题是如何确定你一定要回绝的这部分人所占的比例。首先考虑最明显的策略：回绝最初的一半应聘者，此后，如果出现一个应聘者，其素质比以前的所有应聘者都好，则聘用此应聘者。用这种策略找到最佳人选的概率略大于1/4。

我们分析一下这个策略。考虑两个应聘者在整个队伍中的位置：素质最好的应聘者和素质第二好的应聘者。如果素质第二好的应聘者排在队伍的前二分之一，而素质最好的应聘者排在队伍的后二分之一，则你一定可以得到最佳人选。这个结论非常明显。在其它情况下，你都得不到最佳人选。素质第二好的应聘者排在队伍的前二分之一的概率是 1/2，在这个前提下，而素质最好的应聘者排在队伍的后二分之一的概率略高于 1/2，因为队伍的前二分之一中已经有一个位置被素质第二好的应聘者占据了。所以，你得到最佳人选的概率略大于 25%。

事实上，二分之一并非最佳比例。我们可以把得到最佳人选的概率提高到 37%。具体方法是：回绝前 37% 的应聘者，在后来的应聘者中选一个比先前的应聘者都好的人。下面是解法的细节，有兴趣的读者可以仔细推敲这一段。当你在一个陌生的小镇选择餐馆时，不妨用同样的策略。

如何选择最佳人选

假设应聘者有 N 个，N 不小于 6。你的策略是回绝前 r 个应聘者。以 J 表示素质最好的应聘者。

J 出现在每个位置上的概率相同，都是 1/N。如果 J 出现在前 r 个位置中，你将回绝 J。如果 J 出现在第 r+1 个位置上，你将聘用 J。如果 J 出现在第 r+2 个位置上，你是否聘用 J 取决于前 r+1 个应聘者的排列顺序。在前 r+1 个应聘者中有一个素质最好的人，如果这个人出现在前 r 个位置上，则你将聘用 J——这个概率是 r/（r+1）。如果 J 出现在第 r+3 个位置上，你是否聘用 J 取决于前 r+2 个应聘者的排列顺序。在前 r+2 个应聘者中有一个素质最好的人，如果这个人出现在前 r 个位置上，则你将聘用 J——这个概率是 r/（r+2）。类似地，我们可以得到 J 出现在每个位置上的结果。各个概率值的和即为你得到最佳人选的概率：

143

$$\frac{1}{N}\times\left(1+\frac{r}{r+1}+\frac{r}{r+2}+\ +\frac{r}{r+3}+\cdots+\frac{r}{N-1}\right)。$$

这个值取决于如何选择 r。以 P（r）表示这个值。下面的问题是如何选择 r 才能使 P（r）最大。

以 r+1 代替 r，可以直接得到 P（r+1），即如果回绝前 r+1 个应聘者，你得到最佳人选的概率。比较 P（r）和 P（r+1）的差，如果 P（r）－P（r+1）小于 0，则说明随着最初回绝的人数的增加，得到最佳人选的概率上升；如果 P（r）－P（r+1）大于 0，则结论相反。这个差是

$$P(r)-P(r+1)=\frac{1}{N}\times\left(1-\frac{1}{r+1}-\frac{1}{r+2}-\frac{1}{r+3}-\cdots-\frac{1}{N-1}\right)$$

当 r 较小时，这个值小于 0。当 r 达到某个数之后，这个值始终大于 0。这说明，当 r 较小时，随着 r 的增加概率递增；而当 r 达到某个数之后，概率递减。因此概率最大的条件是 r 是使得 $\frac{r}{r+1}+\frac{r}{r+2}+\frac{r}{r+3}+\cdots+\frac{r}{N-1}>1$ 的最大值。此时

$$P(r)=\frac{r}{N}\times\left(\frac{1}{r}+\frac{1}{r+1}+\frac{1}{r+2}+\frac{1}{r+3}+\cdots+\frac{1}{N-1}\right)\approx\frac{r}{N}$$

根据一个数学中常用的近似公式，此时 r 约等于 N/e，e 是自然对数的底数。因此，P（r）约等于 1/e，近似于 37%。

144

第八章 三局两胜

在两个人（或两支队伍）进行比赛时，通常需要多少局才能决出胜负？在世界杯斯诺克锦标赛中，冠军争夺战曾经采用143局决胜的规定，现在的规则是35局决胜。[1] 在世界杯垒球赛和斯坦雷杯冰球赛中，决赛为七局四胜制。在板球中，由于平局经常出现，比赛所需的局数从两局到六局都很常见。在世界杯国际象棋赛的决战中，比赛的局数经常变化，最常见的是24局决胜。网球赛的规定比较复杂：在每一局中，获胜条件是得分不低于四分，并且领先对手至少两分；在每一节比赛中，获胜条件是至少获胜六局，并且领先对手至少两局；而最终获胜的条件是赢得三节比赛中的两节（或赢得五节比赛中的三节）。在壁球中，双方先争夺发球权，获得发球权的一方可以得分。在每一局中，一方赢满九分时获胜。然而，如果接近终局时打成平局，比分为八比八，则没有发球权的一方有权决定本局的获胜条件：既可以决定得满九分者获胜，也可以决定得满十分者获胜。比赛通常三局决胜或五局决胜。羽毛球的情

① 斯诺克是一种桌球游戏——译者注。

况与壁球类似，但是当获胜条件改变时，规则也要做出调整。这些不同的规则有何优劣？在判断哪一种规则最优时，概率知识是否有用？

多局决胜的目的通常是减少偶然因素对比赛结果的影响，使得比赛的结果能更好地反映双方的真实水平。在桥牌比赛中，如果仅仅凭一手牌定输赢，显然是荒谬的。然而，某些现实因素经常使得多局决胜不可行。在职业拳击赛中，比赛场次过多会严重损伤运动员的身体；在世界杯足球赛中，如果采取多局决胜制，比赛结果反映的可能不是双方的球艺水平，而是身体的耐力。面临实际的抉择时，统计学的理由往往让位于一些更现实的考虑：赛程的长短必须合理，必须保证比赛对参赛者和观众的吸引力。道理很简单：体育运动不是为统计学家设计的。 145

在一些项目中，某一方拥有额外的优势。比如，在国际象棋中先行的一方明显占优势，而在网球赛中，发球方获胜的机会较大。在这样的项目中，比赛双方应当轮换位置。另外一些项目中则没有这个问题，比如足球、摔跤等等。在本章中，我们的目的是研究不同的比赛局数对双方获胜机会的影响。

第一类比赛：斯诺克式比赛

在这一节中，假设我们研究的比赛具备两个特征：

○ 每一局比赛的结果是相互独立的，即某一局比赛的结果对以后的比赛没有影响；

○ 如果在某一局中，某一方获胜的概率是 p，则在以后的一系列比赛中，这一方获胜的概率依然是 p。

如果你想把我们在本节中得出的结论应用于某一场具体的比赛，首先必须检查一下这两个条件是否成立。如果比赛不具备这两个特征，你当然会误入歧途。如果你的研究对象是网球赛，你需要考虑具体情况。假如参赛双方技术平平，那么由谁发球对胜负影响不大，你可以认为这两个条件是成立的；然而，假如比赛在两个职业高手间进行，发球方占有明显的优势，这时应用本节中的结论就是不恰当的。不过，你仍然可以局部地应用本节的结论，比如在具体的某一局中。在赛马中，如果某一个选手在比赛中负伤，此后的比赛结果将受到影响，此时假设的条件不成立。在垒球赛和冰球赛中，如果考虑到主客场对胜负的影响，假设的条件不成立；然而，如果每一局比赛的环境都非常相似，你可以放心地应用本节的结论。

我们研究的核心是一个参赛者在某一局中获胜的概率如何影响其在整个比赛中获胜的概率。假设一个参赛者史密斯在每一局中获胜的概率是 p，比赛规则是先胜 N 局者获胜。由于我们的主要兴趣集中于参赛双方实力接近的情况，所以假设 p 接近于 1/2。用 x + 1/2 表示 p，x 是一个较小的值。

通常，如果某一方已经获胜 N 局，则整场比赛结束。不过，我们可以假定每一场比赛都赛完规定的局数，即使胜负已定，依然把剩下的比赛打完。（世界杯斯诺克锦标赛曾经如此规定过。）这个假定对计算结果没有影响，却可以使我们的研究简化。例如，当比赛规定先胜 5 局者获胜，则我们认为比赛一定进行 9 局。一般地，当获胜条件是先胜 N 局时，我们认为比赛一定进行 2N－1 局。

比赛的结果符合标准的二项分布（关于二项分布的细节请参考附录 2），因为：

○ 比赛的局数固定（2N－1 局）；

○ 每局比赛的结果互不影响；

○ 史密斯在每一局中获胜的概率一定；

○ 决定比赛胜负的因素是史密斯是否赢得至少 N 局。

利用附录 2 介绍的方法，很容易算出在整场比赛中史密斯赢多少局对应的概率。我们把史密斯获胜局数不小于 N 的概率加起来，就得到他赢得整场比赛的概率。

假如比赛规定三局两胜，则 $N=2$。史密斯赢得全部三局的概率是 p^3，恰好赢得两局的概率是 $3p^2(1-p)$，他赢得整场比赛的概率是这两个值的和。用 x 表示 $p-1/2$，则得到最终结果 $\frac{1}{2}+\frac{3x}{2}-2x^3$。当 x 较小时，$x^3$ 可以忽略，所以答案是 $1/2+3x/2$。表 8.1 给出了几个不同的 p值对应的史密斯获胜的概率。当 p 略高于 1/2 时，获胜概率仅比 p 略大一点。这说明当双方实力接近时三局两胜制没有特别突出的优势，然而，当双方实力相差较大时，三局两胜制可以有效地降低强者意外失手的概率。

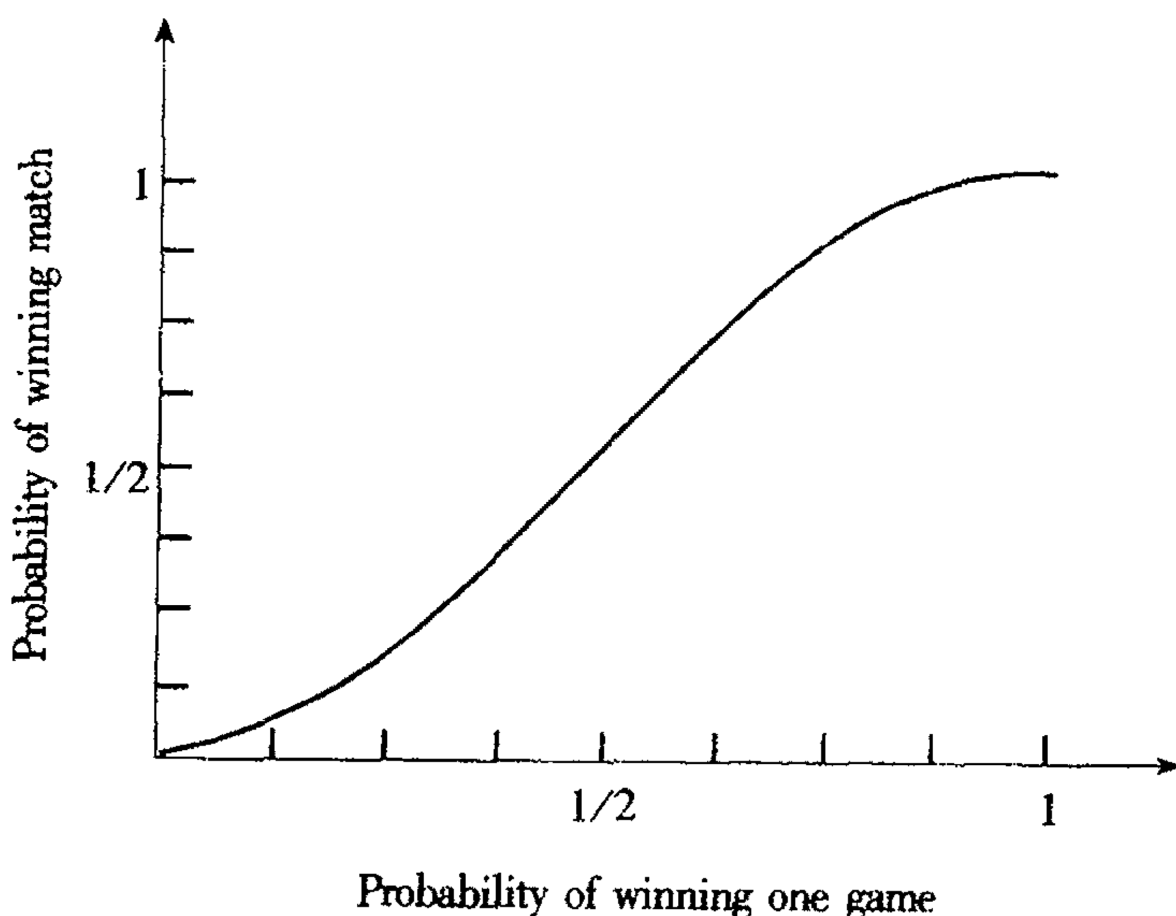
147

表 8.1：在三局两胜的比赛中，
每局的胜率与整场的胜率的关系

每局的胜率	0.52	0.55	0.6	0.67	0.75	0.85
整场的胜率	0.52	0.575	0.643	0.745	0.844	0.939

每局比赛的胜率与整场比赛的胜率的关系也可以用图 8.1 表示。图中横坐标表示每局的胜率，纵坐标表示整场的胜率。

体育比赛中经常规定三局两胜、五局三胜或七局四胜。对于后两种情况，计算方法是一样的。利用附录 2 介绍的方法，很容易得出结论。当史密斯每局的胜率为 60% 时，他在三局



两胜制下的胜率为 65%，在五局三胜制下的胜率为 68%，在七局四胜制下的胜率为 71%。

148 当 N 大于 3 时，可以用同样的方法计算，然而计算过程和结论比较繁琐，本书不作介绍。利用前文介绍过的斯达灵公式，可以得到一个近似公式。当每局的胜率接近于 1/2 时（即 x 接近于 0 时），整场的胜率约等于

$$\frac{1}{2}+x\sqrt{\frac{4N}{\pi}}$$

利用这个公式可以发现决胜局数的增加对在整场的胜率的影响。如果一个人在单独的一局中获胜的概率高于 1/2，则整场获胜的概率也高于 1/2；决胜局数越多，他获得整场胜利的概率也越大。然而，由于 N 出现在开平方的符号下，随着 N 的增加，整场的胜率的增加是非常缓慢的。为了使胜率超过 1/2 的部分增加一倍，N 需要增加到原来的四倍。比如说，史密斯在九局五胜制中的胜率是 55%，为了使他的胜率达到 60%，赛制需要调整为 39 局 20 胜。通常，39 局 20 胜与 29 局

15胜的差别不大。

当双方实力接近时，这个公式相当可靠。表8.2中的几组数据反映了N与整场的胜率的对应关系。假如史密斯每局的胜率为51%，则x=0.01。在19局10胜的比赛中，N=10，史密斯最终获胜的概率是53.5%。然而，假如假设史密斯每局的胜率为60%，则x=0.1。在19局10胜的比赛中，史密斯最终获胜的概率是85%。在7局4胜的比赛中，N=4。如果一支队伍每局的胜率为55%，则x=0.05，其最终获胜的概率接近于61%。如果这支队伍每局的胜率为60%，则最终获胜的概率约为71%。如果这支队伍每局的胜率仅仅略大于50%，为了使他们最终获胜的概率明显高于50%，需要使N变得非常非常大。

表8.2：N与整场的胜率的对应关系

149

N	1	2	3	4	5	7	10	16
整场胜率 - 1/2	x	1.5x	1.875x	2.1875x	2.461x	2.933x	3.524x	4.478x

表8.3：N和p（每局的胜率）决定整场的胜率

N	5	10	15	20	25	30
P=0.52	0.55	0.57	0.59	0.60	0.61	0.62
P=0.55	0.62	0.67	0.71	0.74	0.76	0.78
P=0.60	0.73	0.81	0.86	0.90	0.92	0.94

应用这个公式时应当小心。对于确定的x的值，当N足够大时，公式的计算结果有可能超过1。概率大于1显然是荒唐的。因此，对于任何的x的值，N都不能超过一定范围。在什

么范围内这个公式可以安全地使用呢？只要保证 N 小于 $1/(10x^2)$就可以了。

为了使实力较强的选手在比赛中拥有明显的优势，比赛至少应当进行多少局？表 8.3 给出了部分答案。在世界杯斯诺克赛中，实力较强的选手作为种子选手参赛，首轮比赛为 19 局 10 胜制。如果种子选手的每局胜率接近 60%，则他获得整场胜利的概率接近 80%。在以后的各轮比赛中，比赛局数递增。半决赛为 31 局 16 胜，决赛为 35 局 18 胜。在最后阶段，我们可以假设参赛选手的实力很接近，通常 p 在 45%到 55%之间。此时实力较强的选手仍然有可能失手，不过概率只有 1/4。

在严肃的比赛中，对局数是事先规定的。然而，如果比赛在朋友之间进行，规则往往可以变通。假如古罗马的尼禄皇帝参加斯诺克赛，他将按照自己的意愿不断调整规则。尼禄的规则如下：假如第一局尼禄获胜，则规则为一局决胜，大家庆祝皇帝伟大的胜利；如果第一局对手获胜，则规则调整为三局两胜；如果对手再次获胜，则规则调整为五局三胜；以此类推。一旦尼禄获胜的局数领先于对手，则比赛结束，尼禄获胜。在这种规则下，尼禄最终获胜的概率是多少？

这个问题应当分三种情况逐一讨论。

情况一：尼禄的每局胜率 p 大于 50%；

情况二：尼禄的每局胜率 p 等于 50%；

情况三：尼禄的每局胜率 p 小于 50%；

在情况一中，答案很明显。当比赛进行了很多轮时，尼禄获胜的局数超过 1/2，尼禄必然获胜。在情况二中，当比赛进行了很多轮时，尼禄获胜的局数接近于 1/2，有时候对手获胜局数领先，有时候尼禄获胜局数领先，但是对手的获胜局数始终领先的概率接近于 0。所以，无论在情况一中还是在情况二中，尼禄最

终获胜的概率都是 1。

情况三则非常有趣。当比赛进行了很多轮时，尼禄获胜的局数所占的比例接近于 p，所以从长期看，尼禄获胜的局数落后于对手，并且始终保持落后。然而，在比赛的初期，尼禄有可能领先。当 p 小于 1/2 时，尼禄领先的概率（也就是尼禄最终获胜的概率）为 p/(1-p)。例如，当 P=1/3 时，尼禄最终获胜的概率为 1/2。当 p 略低于 1/2 时，尼禄获胜的概率相当高。如果 p=45%，则这个概率为 45/55，约等于 89%。但是当 p 仅为 10% 时，尼禄的机会只有 11%。下面是这个问题的解法的细节。

尼禄领先的概率

用 X 表示尼禄领先的概率，用 p 表示尼禄在每一局中获胜的概率。第一局结束之后，有两种可能结果：

1. 尼禄获胜。此时尼禄领先。这个结果出现的概率为 p；
2. 对手获胜。此时尼禄落后 1 局。这个结果出现的概率为 1-p。

假设第一局对手获胜。尼禄为了最终领先，必须完成两个目标：

A. 把总比分追平；

B. 在把比分追平之后，再超出 1 局。

为了完成目标 A，尼禄必须比对手多赢 1 局，也就是说，尼禄必须在这个阶段中领先。因此，尼禄实现目标 A 的概率就是 X。在实现目标 A 的前提下，尼禄实现目标 B 的概率也是 X。因此，尼禄同时完成这两个目标的概率是 X^2。所以，

$$X = p + (1-P)X^2$$

解这个方程，可以得到两个解：X=1 以及 X=p/(1-p)。由于 X 是一个概率，所以 X 不可能大于 1。当 p 大于 1/2 时，p/(1-p) 大于 1，所以答案只能是另一个解，X=1；当 p 等于 1/2 时，p/(1-p)=1，两个解相等，所以答案依然是 X=1。

当 p 小于 1/2 时，两个解都可能是最后答案。然而，我们不能接受"尼禄必然获胜"这个结论。所以，答案是 X=p/(1-p)。

151

第二类比赛：壁球式比赛

传统的壁球比赛规则是只有发球的一方可以得分。虽然现在的许多壁球比赛规定，发球方和接球方都可以得分，我们在本节中只考虑传统规则下的情况。首先我们要考虑两个概率：一个选手在一轮争夺中赢得一分的概率（假设他是发球方）以及这个选手在一轮争夺中赢得发球权的概率（假设他是接球方）。我们假定这两个概率是相等的，用 p 表示。此外我们考虑两个概率：当你是发球方时，在下一轮中你得分的概率 S；以及当你是接球方时，在下一轮中你得分的概率 R。S 和 R 都可以由 p 计算出来，下面是计算的细节，没兴趣的读者不必深究这一段，记住结论就可以了。

图 8.2 表示 S 和 R 与 p 的关系。请注意，当 p 接近于 1/2 时，S 和 R 相差很大。这可以解释为什么在两个实力非常接近的选手之间也经常出现悬殊的比分。当 $p = 1/2$ 时，$S = 2/3$，这说明当双方势均力敌时，发球方连续得分的机会很大。

如果现在场上比分为 8:8，你是接球方。你是否应当要求 152 把获胜条件调整为先得 10 分者获胜？为了做出决定，我们计算一下在两种情况下你的获胜机会。如果你不要求调整获胜条件，则先得 9 分者获胜。为了获胜，你必须赢得下一分。由于你是接球方，这个概率是 R。如果你要求把获胜条件调整为先得 10 者获胜，你有三种获胜途径：

1. 连赢两分；
2. 赢一分，输一分，再赢一分；
3. 输一分，然后连赢两分。 153

S和R与p的关系

当你是接球方时，为了得分你必须先夺得发球权。你在下一轮中夺得发球权的概率是p，而夺得发球权后你就变为发球方，所以

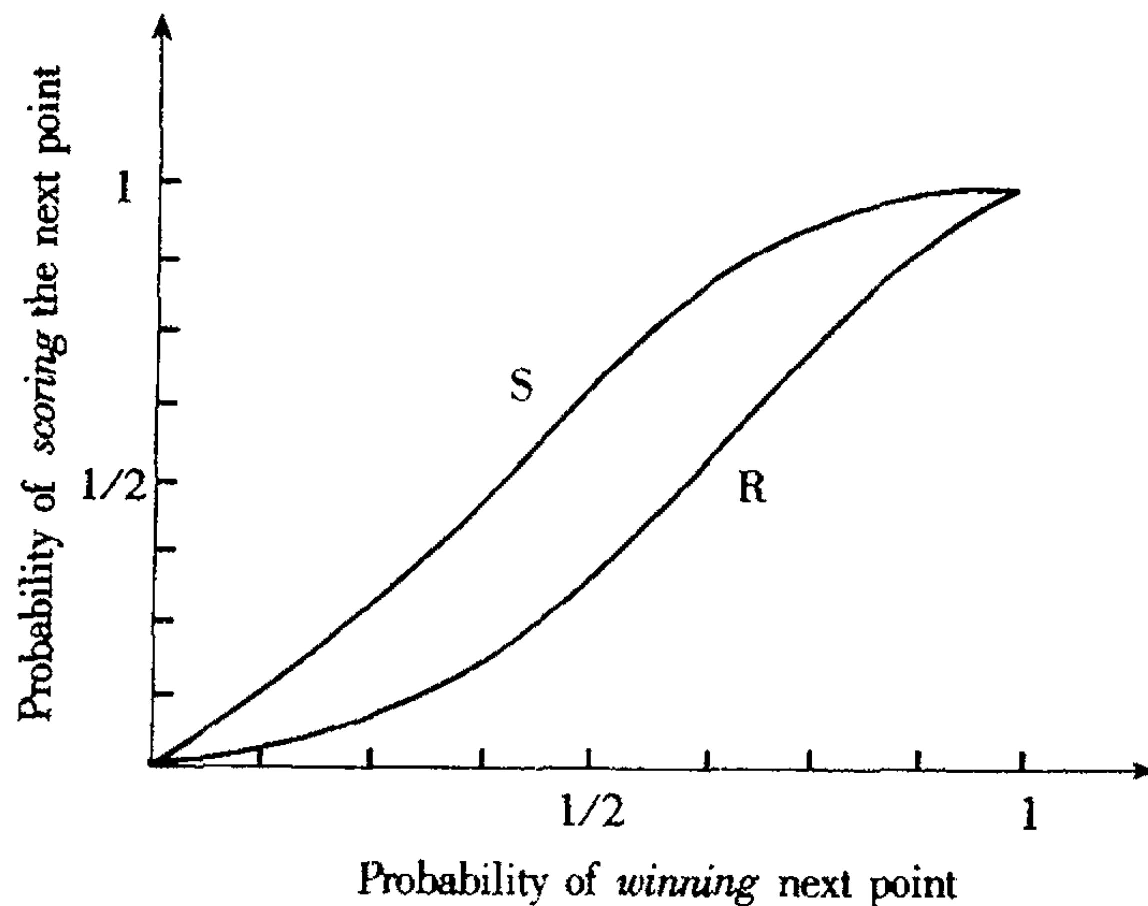
$$R = pS$$

当你是发球方时，你有两种得分的途径：其一，在下一轮中直接得分，其概率为p；其二，在下一轮中丢掉发球权（概率为1－p)，而作为接球方得分。所以

$$S = p + (1-p)R$$

把两个式子综合起来，就得到结论：

$$S = p/(1-p+p^2),\ R = p^2/(1-p+p^2)$$



途径1的概率是R×S（赢第一分时你是接球方，赢第二分时你是发球方）；途径2的概率是R×（1－S）×R（赢第一分时你是接球方，输第二分时你是发球方，赢第三分时你是接球方）；途径3的概率是（1－R）×R×S。这三个概率的和是你在调整获胜条件之后的获胜概率。

哪种情况对你更有利？答案是当S大于1/2时，后者较好。S的值是由p决定的，当p大于38%时，S大于1/2。所以结论是：

当p大于38%时，应当要求把获胜条件调整为先得10分者获胜。在实战中，这个条件几乎总是被满足。因为你们的比分已经达到8:8，说明双方实力接近，p应当在50%附近，通常高于38%。当然，有时会出现例外。比如你的对手体力充沛而你已经感到体力不支，或者你刚刚受伤，此时速战速决是上策。总之，当p小于38%时，应当保持获胜条件不变，而当p大于38%时，应当调整获胜条件。

下面我们讨论赢得每一分的概率如何影响赢得整个一局比赛的概率。相信你已经发现，有两种情况都能使比分达到2比1：你赢1分、输1分、再赢1分以及先输1分、然后连赢2分，但是这两种情况对应的概率不同。如果我们专注于具体的比分上升的过程，则问题变得极其复杂，许许多多的情况需要分别考虑。这种策略太麻烦了，我们应当找一种方便的算法。

为了简化我们的讨论，假定作为发球方和作为接球方得分的概率是相同的。当两个选手实力相当时，假定两种胜利条件（即先得9分者胜及先得10分者胜）被应用的频率相同。为了取胜，你必须作为发球方获胜9或10个回合；为了赢得发球权，你还必须作为接球方获胜若干回合，这个数大致也是9或10。因此，这个问题转化为一场没有发球权问题的普通比赛，

获胜条件为赢得18~20轮。当你的实力比对手强时，你控制发球权的时间比对手长，所以你需要的获胜轮次比18~20要小一些，比如16~18。（假如比赛规则规定发球方和接球方都可以得分，则获胜条件通常是赢得15分。）

154

当p在40%到60%之间时，问题可以转化为斯诺克式的比赛，我们可以直接利用上一节中的结论，N的取值为16~18。利用表8.2和8.3可以直接得出近似结果。我用计算机模拟了N的取值范围为16~18、p的取值范围为45%~55%的情况，结论支持这个近似算法。当然，这种计算方法是有缺陷的，但可以提供非常好的近似结果。由于我们最后算出的概率与N的平方根有关，而当N的取值范围为16~18时N的平方根很小，所以这种方法的误差很小。我们用一个例子说明计算过程：当p=51%时，在表8.2中可以找到N=16所对应的结果，你赢得一局的概率为55%。再进一步，如果整个比赛为三局两胜，再次利用表8.2可以得到你赢得整场比赛的概率，p=55%对应的结论是57.5%。如果整个比赛为五局三胜，表8.2给出的结论是你有59%的机会赢得整场比赛。在附录2中有计算公式。

第三类比赛：网球式比赛

网球比赛的特点是，无论整场比赛已经进行了多久，得分较少的一方有可能取胜。在某一节比赛中也是如此：获胜者有可能是得分较少的一方。比如说，你在一节比赛中拿下六局，而每局都仅仅领先两分；对手拿下了四局，而每局都把你打成0蛋。这样，他的得分比你高，但胜利者是你。

G与p的关系

首先考虑比赛延长时的情况。当双方打成3:3时，比赛将延长，因为获胜方必须超出对手两分。用D表示史密斯在这种条件下获胜的概率。如果史密斯连得两分（概率为p^2），则史密斯获胜；如果对手连得两分（概率为q^2），则对手获胜；如果双方各得一分（概率为2pq），再次出现平局，回到最初的情况。因此

$$D = p^2 + 2pqD。$$

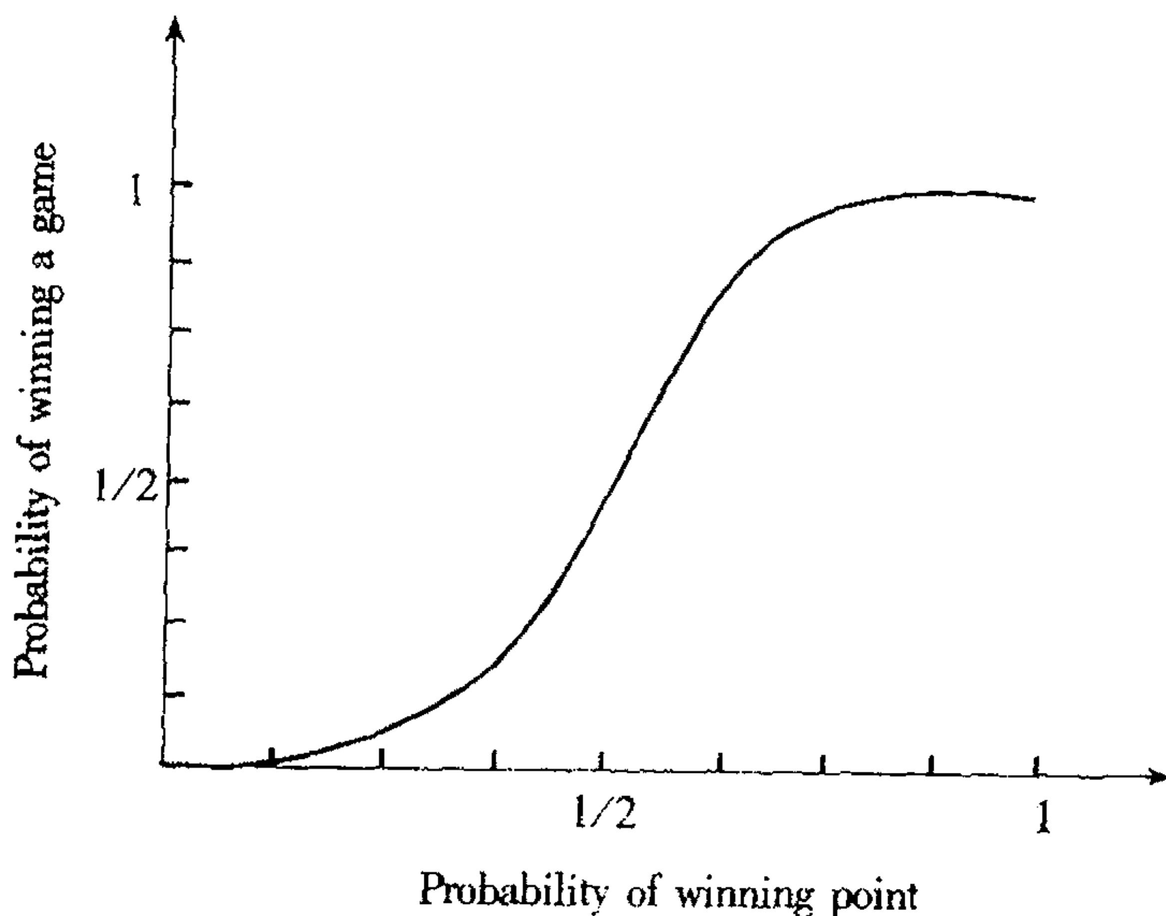
即$D = p^2/(p^2 + q^2)$。

史密斯可以在以下四种条件下获胜：

1. 连得4分，概率为p^4；
2. 在前4分中得到3分，而后得到自己的第4分，概率为$4p^4q$；
3. 在前5分中得到3分，而后得到自己的第4分，概率为$10p^4q^2$；
4. 在前6分中得到3分，而后在比赛延长的情况下获胜，概率为$20p^3p^3D$。

把这四个值相加，得到

$$G = (p^4 - 16p^4q^4)/(p^4 - q^4)$$



在每一局的内部，获胜条件是至少领先两分，得分不低于4。由于没有交换发球权的问题，所以可以直接引用斯诺克式比赛中的结论，只要把“局”换成“分”就可以了。用 p 表示史密斯赢得1分的概率，用 q 表示 $1-p$，q 即史密斯丢1分的概率。当 $p=1/2$ 时，史密斯最终获胜的概率显然是1/2，所以，我们可以把注意力集中于 p 不等于1/2的情况。用 G 表示史密斯赢得一局的概率。下面是计算 G 的过程。

图8.3展现了 G 与 p 的关系。随着 p 的增加，G 递增。当 p 刚刚超过1/2时，G 迅速上升。当 p 达到70%时，G 达到90%。而此后，随着 p 的增加 G 的上升非常缓慢。这说明，假如你在发球局中得分的概率已达到70%，则进一步加强你在发球局中的优势对你的好处不是特别大，因为你的胜率已经高达90%，这个数不容易变得更大。此时你应致力于提高自己在接球局中的胜率。1997年温布尔登公开赛中的统计数据支持我们的结论。赫曼在发球局中的得分率为75%，最终赢得23个发球局中的22个。克拉克在发球局中的得分率为73%，最终赢得23个发球局中的21个。在女子比赛中，辛吉斯在发球局中的得分率为55%，她赢得13个发球局中的9个。诺娜在发球局中的得分率为54%，她赢得13个发球局中的8个。

当 p 接近于1/2时，可以立刻得到一个关于 G 的近似公式。把 p 表示为 $1/2+x$，则 x 接近于0。G 可以表示为

$$G=1/2+5x/2-\cdots\cdots$$

当 x 接近于0时，公式中被省略的部分非常小，可以忽略不计。其中 x 的系数5/2说明每局胜率增长的速度比得分率增长的速度快。例如，当你的得分率从50%增加到52%时，你的胜率从50%增加到55%。

以上我们讨论的是赢得1分的概率如何影响赢得1局的概

率。下面我们还要研究赢得1局的概率如何影响赢得1节的概率。在1节比赛中，双方的发球局与接球局交替，获胜条件是拿下6局并超出对手2局。在发球局和接球局中，一个选手获胜的概率不同。假设有一个选手琼斯，他在发球局中的胜率为s，而在接球局中的胜率为r。根据1997年温布尔登公开赛中的统计数据，s通常大于90%，而r通常小于10%。

在网球比赛规则引进“最后决胜局”以前，1节比赛往往进行很多局，可能出现15:13或24:22之类的大比分。有些比赛规定，如果某一节比赛是全场比赛的最后一节，则在这节比赛中没有最后决胜局。在没有最后决胜局的情况下，计算琼斯赢得本节比赛的概率的方法与以前的算法类似，但我们需要考虑的是两个概率值（r和s），这使得解题过程变得极为繁琐。157 我们略过计算的细节，直接给出结论。

决定琼斯在1节比赛中的胜率的关键因素是r和s的平均值。用P表示（r+s）/2。当P=1/2时，琼斯获胜的概率为50%。如果希望获胜的概率超过50%，则应当使P大于1/2。网球教练都知道一个基本原则：弥补自己的短处通常比加强自己的长处更有效。如果s已经达到90%，无论你如何努力也不能把它变成100%；相反，如果r只有10%，则你还有很多可以努力的空间。把10%提高到15%要比把90%提高到95%容易得多，但是从结果看，二者没什么差别。

在最后决胜局中，一名选手先发一次球，其后双方轮换发球，每轮连发两次球。获胜条件是拿下7分，并且至少领先对手2分。计算在最后决胜局中的胜率与计算在一节中的胜率非常相似，只不过6变成了7。用s表示琼斯在发球时得分的概率，用r表示琼斯在接球时得分的概率，关键性的因素依然是（r+s）/2。通过1997年美国公开赛的统计数字，我们可以发

现顶尖高手之间的实力相当接近。在一场半决赛中，每节的比分为6:3、6:3、6:4，双方的总得分为97:78，失败者的得分率达到了45%。在另一场半决赛中，每节的比分为6:1、3:6、3:6、6:3、7:5，双方的总得分为128:127。表8.4和表8.5给出了几个典型的s值和r值对应的结论。表8.4表示最后决胜局中的情况，第一列表示不同的s值，第一行表示不同的r值，表中的元素表示此选手赢得最后决胜局的概率。表8.5表示在没有最后决胜局的条件下整节比赛的情况，第一列表示不同的s值，第一行表示不同的r值，表中的元素表示在没有最后决胜局的情况下此选手赢得一节比赛的概率。（请注意，在表8.5中s和r分别表示的是此选手在接球局和发球局中的胜率。）

表8.4：在最后决胜局中获胜的概率

	0.5	0.4	0.3	0.2
0.5	0.500	0.345	0.206	0.098
0.6	0.655	0.500	0.337	0.185
0.7	0.794	0.663	0.500	0.317
0.8	0.920	0.815	0.683	0.500

158

制定规则的目的就是使比赛公平。在最后决胜局中，当双方的得分之和为奇数时，显然有一方比对手多发了一次球，然而，关于发球次序的规定使得双方轮流持有这个优势。关于最后决胜局的规定和关于一节比赛的程序的规定都使得双方轮换发球权，这使得比赛对于双方是公平的。

表 8.5：在没有决胜局的条件下赢得整节比赛的概率

	0.5	0.4	0.3	0.2	0.1
0.5	0.500	0.354	0.221	0.113	0.038
0.6	0.646	0.500	0.347	0.200	0.079
0.7	0.779	0.653	0.500	0.327	0.150
0.8	0.887	0.800	0.673	0.500	0.275
0.9	0.962	0.921	0.850	0.725	0.500

在最后决胜局中以及在一节比赛中都有轮换发球权的问题，但是两种轮换的方式不同。根据我们的数学模型，两种轮换方式的结果是一样的。无论采取哪种轮换方式，一名选手顶多比对手多得一次发球权，然而获胜需要领先两分（或两局）。不过仍然有些球迷认为，没有最后决胜局的比赛规则是不公平的，在一节比赛中，先发球的一方有利。其理由是，假设你先发球，而整节比赛进行到 4:4，如果你在这个阶段丢掉一个发球局，你仍然有机会追回；相反，假设你后发球，而整节比赛进行到 4:4，如果你在这个阶段丢掉一个发球局，你就死定了。即使这种观点有道理，我们也可以用一个简单的规定把规则调整到公平的程度：

在一节比赛中，执行在最后决胜局使用的轮换发球权的规定。

小　结

一场比赛不能占用太长的时间。在实际情况中，通常双方

进行若干局比赛，先拿下 N 局者获胜。我们可以计算出每局比赛的胜率如何影响整场比赛的胜率。我们的计算依赖于一个重要的假定：每局比赛的结果互不干扰。我们假定参赛双方在比赛中始终保持平和的心态，不受比赛成绩的影响。

当双方实力接近时，三局两胜制与五局三胜制差别不大。由于比赛进行的局数很少，这种规则的主要意义只是降低强者意外失手的风险。如果一方的每局胜率达到 60%，则当 N 达到 10 时就可以保证这一方的胜率不低于 80%；然而，如果一方的胜率仅仅略高于 50%，为了使这一方最终的胜率达到 80%，必须使 N 非常大。整场胜率的增长与 N 的平方根大致成正比。

159

另外两个模型研究有发球权的情况。壁球式比赛代表以发球权为得分的前提的情况，网球式比赛代表发球方得分率较高的情况。

160

第九章　电视游戏

电视游戏在全球范围内流行。有些电视游戏考验参与者的知识和技能，另外一些电视游戏则完全凭借运气。本章将介绍四种电视游戏，并讨论概率知识在这些游戏中的应用。

挑盒子

挑盒子是一个相当著名的游戏。有三个完全相同的盒子，分别标为 A、B、C。在其中一个盒子中，放有一笔奖金，另外两个盒子是空的。游戏的主持人哈福知道奖金在哪个盒子里。游戏的参与者玛丽可以在三个盒子里挑一个，她的目标是挑中那个装有奖金的盒子。假设玛丽挑了盒子 B。

不论玛丽挑了哪个盒子，在剩下的两个盒子里至少有一个盒子是空的。在玛丽做出决定之后，哈福打开一个空盒子给玛丽看，并告诉玛丽，她现在有权利改变自己的决定，选择另外一个盒子。几乎在所有情况下，玛丽都会拒绝改变决定。玛丽的逻辑是，无论她一开始选择哪个盒子，哈福都可以在剩下的

盒子中拿出一个空盒子给她看，所以哈福的行为对她最初的选择的正确与否没有任何影响。但是，玛丽的想法是错误的。玛丽确实应当改变决定。在此情况下，**玛丽改变决定可以使成功的概率提高一倍**。

我知道很多人反对这个结论，不过请继续往下读。有一个名叫玛丽莲的女士为一份美国杂志《帕拉得》主持读者问答专栏（据说此女士是世界上智商最高的人），她的专栏介绍了这个问题的正确答案。在此之后，有很多人写信批评她的结论，而且有些批评出自专业的统计学教授之手。然而，玛丽莲的观点是正确的，她的分析惊人地简单。

我们首先研究最简单的情况：如果玛丽坚持最初的决定，她的选择正确的概率是多少？由于她在三个盒子中随机地选择一个，所以她的选择正确的概率是 1/3。

161

玛丽最初选择空盒子的概率是 2/3。这种情况发生时，在剩下的两个盒子中只有一个是空盒子。此时主持人只有一个选择，即打开剩下的空盒子。于是，最后的一个盒子就是装有奖金的盒子。如果玛丽此时改变最初的决定，她就可以获得成功。我们已经知道，这种情况发生的概率是 2/3，所以如果玛丽改变最初的决定，她成功的概率是 2/3。她成功的概率增加了一倍。

我曾劝说一个电视台实际检验这个理论。这个电视台在一个中学里找到 53 个孩子参与这个游戏。每个孩子参与两次，一次坚持最初的决定，另一次改变最初的决定。根据理论预测，平均第一种策略将成功 17.67 次，而第二种策略将成功 35.33 次。实际结果如何呢？第一种策略成功了 19 次，而第二种策略成功了 36 次。多么精确的符合！

正如在大英帝国彩票中有些人始终不相信 {1，2，3，4，

5，6｝是一个出现频率极高的号码组合，在这个问题中许多人始终不相信我的结论。如果你觉得我的结论不合理，不妨考虑这种情况：主持人提供100个盒子，其中99个是空的，只有一个盒子中有奖金。你选择一个盒子，而后主持人打开98个盒子。此时你可以坚持最初的选择，也可以改变选择。你是否应当改变选择？

如果你拒绝改变，你只有在一开始就选择了正确的盒子的情况下才能获得成功，这个概率只有1%。在另外的99%的情况下，你最初选择的是一个空盒子，而另外的98个空盒子已经打开，你这样改变最初的选择就可以成功。所以，在99%的概率下，改变选择是正确的。

如果你能接受这个例子，为什么你不能接受最初的例子呢？

下面是一个类似的问题。一个家庭中有两个孩子：老大和老二。每个孩子是男孩或女孩的概率相同。显然有三种可能：两个孩子都是男孩；一个孩子是男孩而另一个是女孩；两个孩子都是女孩。三种情况的概率各为多少？相信你可以解决这个简单的问题。答案是1/4、1/2、1/4。

如果我给你一个额外信息："老大是男孩"，结果如何呢？由于老大的性别已确定，老二是男孩或女孩的概率都是1/2，所以只有两种可能：两个孩子都是男孩或者一个孩子是男孩而另一个是女孩，概率都是1/2。

还是这个问题。如果我给你的额外信息变成："老二是男孩"，结果如何呢？由于老二的性别已确定，老大是男孩或女孩的概率都是1/2，所以只有两种可能：两个孩子都是男孩或者一个孩子是男孩而另一个是女孩，概率都是1/2。

请注意，最玄妙的地方到了。假如我给你的额外信息是

"两个孩子中有一个是男孩"，结果如何呢？

你是否同意这种观点："有两种可能：其一，老大是男孩；其二，老二是男孩。无论在哪种情况下，答案都是一样的：两个孩子都是男孩或者一个孩子是男孩而另一个是女孩，概率都是1/2。所以这也是最后的答案。"这个结论是错误的！正确的答案是两个孩子都是男孩的概率为1/3，一个孩子是男孩而另一个是女孩的概率是2/3。理由如下：最初有四种可能，（1）两个孩子都是男孩；（2）老大是男孩而老二是女孩；（3）老二是男孩而老大是女孩；（4）两个孩子都是女孩。四种情况的概率相同。在你得到额外信息"两个孩子中有一个是男孩"之后，最后一种可能被排除，剩下的三种可能概率相同。

同挑盒子的问题一样，这个问题的关键在于额外信息对最初的概率分布有何影响。在挑盒子的问题中，如果玛丽最初挑中的是正确的盒子，则主持人可以在剩下的两个盒子中任取一个打开；而如果玛丽最初挑中的不是正确的盒子，则主持人的选择是惟一的。同样，在这个问题中，如果额外信息是"老大是男孩"，则你可以排除（3）和（4）两种可能；如果额外信息是"老二是男孩"，则你可以排除（2）和（4）两种可能；但是如果额外信息是"两个孩子中有一个是男孩"，你只能排除（4）一种可能。

猜猜看

这个游戏由三个人参与。游戏分三个阶段：第一个阶段淘汰一名参与者，这个阶段进行的时间较长；第二个阶段再淘汰一名参与者，游戏以投标的方式开始；在第三个阶段中，最后

剩下的一名参与者有机会赢得一笔奖金。

三个阶段的规则略有不同，但主题是相同的。主持人向参与者提供一些对象，并明确告知其中有若干个对象具有相同的性质。参与者的任务是把这些具有相同性质的对象挑出来。有时候参与者要应用自己的知识，有时候则完全依靠猜测。下面这个例子发生在游戏的第二阶段，参与者是朱丽和娜塔莎。主持人提供 12 个人名，其中恰好有 8 个是狄更斯小说中的人物。这 12 个人名排列如下：①

A	B	C	D
E	F	G	H
I	J	K	L

游戏最初，两个参与者要进行“投标”，即承诺自己可以挑出多少个正确答案。当然，投标额必须在 1 至 8 之间。下面是游戏进程的描述，括号中的内容是参与者的内心独白，引号中的内容是参与者的话。

朱丽：（天哪，我对狄更斯一无所知！我只知道 A 肯定是正确答案，I 肯定不是，其他的我只好瞎蒙了。如果我的投标额是 2，观众一定会笑翻了。我必须显得自信些。）“4 个。”

娜塔莎：（我可以确定 4 个正确答案，而且我知道 E 肯定不是。朱丽肯定知道 4 个正确答案，我得比她多。）“5 个。”

朱丽：（哎呀！娜塔莎真厉害！打死我我也说不出 6 个。

① 原文中给出的是英文人名。具体人名与主题无关，所以下面仅以英文字母代替——译者注。

就让她先说吧。）“同意，5个。”

娜塔莎：“A。”——观众鼓掌。

“G。”——热烈的掌声。

“L。”——观众欢呼。

“B。”——观众再次欢呼。

“H。”——观众安静下来。

主持人：“非常抱歉。H不是。

娜塔莎没能完成任务，现在轮到朱丽。现在朱丽只要说出一个正确答案就可以获胜。如果朱丽也失败，则娜塔莎又得到一次机会。只要有一个人再说出一个正确答案，游戏即告一段落，说出最后一个答案的人获胜。

朱丽：（碰碰运气吧。）“F。”——掌声四起。

虽然娜塔莎的知识比朱丽丰富，但最后的胜利者是朱丽。朱丽的心理战很成功，她把对手推上了一个很高的目标。 164

在这个过程中，数学知识有什么用呢？数学的用处是帮你确定投标额。你的目的在于：要么投标一个较小的数额，而自己先回答；要么迫使对手投标一个较大的数额。如果你中标，投标额越小越好，这样你获胜的概率较大；如果对手中标，投标额越大越好，这样对手犯错误的概率较大。在你们两人的知识范围相同的情况下，最佳的投标策略是什么？答案很简单：在确保自己完成任务的概率不低于50%的前提下，投标额尽量高。

下面解释一下这个策略。我们比较所有可能的策略。假设B是保证你完成任务的概率不低于50%的最高投标额。如果你最初的投标额是B，而对手不再加码，则你完成任务的概率不低于50%，即你获胜的概率不低于50%，情况显然对你有利；如果你最初的投标额低于B，对手就会选择投标额B。由于你

们两个人的知识范围相同，对手完成任务的概率也不低于50%，你获胜的概率不超过50%，情况对你不利；最后，如果你的投标额高于B，则对手不加码，你先回答问题，则你完成任务的概率低于50%，情况对你不利。综上所述，B是最佳投标额。

请仔细推敲以上推理。当你面对一个问题不知所措时，不妨把自己放到对手的位置上，从对手的角度分析问题。考虑一下对于你的各种战略对手可能做出什么反应，往往可以帮助你找到最佳策略。在两个人对抗的游戏中，这通常是发现最佳策略的最有效的方法。

在这个例子中，我们如何确定B的值呢？B的值取决于你的知识范围。在所有候选答案中，有一些你可以肯定是正确答案，有一些你可以肯定不是正确答案，剩下的就要靠猜测。我们从这种情况入手：假定你确知其中的4个是正确答案，1个不是正确答案，而剩下的7个你不能确定。你只知道这7个候选答案中有4个是正确答案，3个不是正确答案。当你不得不在这7个不确定的答案中做出选择时，假定你的选择是完全随机的，其中的任何一个被选中的概率是相同的。

在此情况下，如果你的任务是找出4个正确答案，你肯定可以完成。考虑你的任务是找出5个正确答案的情况。你首先可以挑出4个确切的答案，然后从7个不确定的答案中随机地猜1个。其中有4个正确答案，所以你有4/7的概率一举获胜，另外有3/7的概率猜错。不过，在你猜错以后，你可能还有下一次机会。如果你猜错了，轮到你的对手回答问题。现在的情况是已经有4个正确答案被挑出来，并且有一个错误答案已经被标明。还剩下7个候选答案，其中有4个正确，3个不正确。在3个不正确的答案中，有一个是你知道的。你的对手

只要猜对一个正确答案即可获胜。

由于对手的知识范围与你相同（这是我们的初始假定），所以对手也知道3个不正确的答案中的一个。因此，对手所要做的就是在6个不确定的答案中随机地选一个，其中4个是正确的，2个是不正确的。对手成功的概率是4/6，失败的概率是2/6。如果对手失败，你就可以获得第二次机会。你获得第二次机会的概率是（3/7）×（2/6）。在你获得第二次机会的情况下，还有6个不确定的答案剩下来，其中有4个正确的答案，2个不正确的答案，而在不正确的答案中，有一个是你知道的。所以你的任务就是在5个不确定的答案中选1个，5个答案中有4个是正确的，你成功的概率是4/5。总之，你通过第二次机会获胜的概率为 165

$$(3/7) \times (2/6) \times (4/5) = 4/35$$

如果你在第二次机会中失败，下一轮对手必然获得成功，因此你不会再得到第三次机会。综上所述，如果你的投标额是5，并且由你先回答问题，你最终获胜的概率为 $4/7 + 4/35 = 24/35$。

你是否可以投标6呢？此时你有三种获胜的途径：1. 一次回答出6个正确答案；2. 第一次只回答出4个正确答案，第5个答案猜错（这是你的第一次猜测），然后由对手继续，但最终仍然是你获胜；3. 第一次回答出5个正确答案，包括1个猜测的结果，然后由对手继续，但最终仍然是你获胜。读者可以独立计算这三种情况对应的概率，我直接给出结论：三个概率值的总和为17/35，这说明你获胜的概率低于50%。所以，你不应投标6。

如果你投标5，对手应当怎么办呢？如果对手的数学知识和你一样渊博，他会发现：假如他让你先回答，而获胜条件为

挑出5个正确答案，你的胜率为24/35，于是他的机会只有1 − 24/35 = 11/35；此时如果他投标6，他可以把先回答问题的权利夺过来，而获胜的条件为挑出6个正确答案。此时他的胜率为17/35，显然比11/35好。所以，如果你们两个都按最佳策略行事，游戏过程应当是这样：

你：（我知道A、D、K和L是正确答案，I不是。剩下的我只好瞎蒙了。如果我投标4，对手一定不会给我先回答的权利。如果我投标6，对手会让我先回答。但是我必须猜两次，两次都猜对的机会不大，我获胜的概率低于50%。我应当显得很不自信，这样对手可能掉以轻心，让我先回答。）"我想一想……5个吧。"

对手：（我知道4个正确答案和1个不正确答案。我估计对方和我知道的一样多。如果我让对方先回答，而获胜条件是挑出5个正确答案，对方的机会很好。本来我想投标5，可惜让对方占先了。如果我投标6，我获胜的机会不大，但对我来说这已经是最好的选择了。我应当显得自信一点，这样对方很可能投标7，我就通过心理战术获利了。）"6个。"

你：（好极了。我不会投标7——傻瓜才会那样干。）"同意，6个。"①

以上结论依赖于一些特定的假设。如果你希望在实战中应用这种分析方法，你实际面对的问题可能与这个问题有出入。

---

① 此处作者的分析有误。在这个例子中，如果你投标6并且先回答，你可以使自己获胜的概率达到18/35。具体策略如下：首先回答4个正确答案，而在第5次，有意回答1个错误答案。此时对手必须率先开始猜测，问题转化为对手投标6而且先回答的情况。根据作者的分析，对手获胜的概率为17/35，所以你获胜的概率为18/35。这说明，在此例中投标6优于投标5。在数学上作者是正确的，但作者忽略了这个诡计。——译者注

你必须根据实际情况做出调整。比如说，你们两个都知道4个正确答案和1个不正确答案，但是你们所知道的内容并不完全重合——这时最佳策略会不会有所不同？再比如，你在不确定的答案中的选择不是完全随机的，你认为E有70%的可能是正确答案，而G有80%的可能不是正确答案——这时问题变得更加复杂。 166

假设你们两个都知道4个正确答案和1个不正确答案，但是你们所知道的内容并不完全重合。不妨假定你们两个都知道G和L是正确答案，此外你知道D和J是正确答案，而对手知道A和H是正确答案。在此情况下，如果你投标5而对手让你先回答，你必须一举成功。如果你第一次失败了，对手不会给你第二次机会。你可以先说出4个你确知的答案，然后瞎蒙1个，你猜对的概率是4/7。假如你不幸猜错了，对手只要在剩下的答案中挑出1个正确答案就可以成功，这对于他当然不成问题。所以，投标5的胜率是4/7，高于50%。如果你投标6而对手让你先回答，你必须瞎蒙两次，两次都猜对的概率只有（4/7）×（3/6）=2/7，而一旦你猜错，对手不会再给你机会。所以，在此情况下投标5仍然是最佳策略。

对于每个不确定的答案，如果你知道在一定概率上其为正确答案，情况如何呢？此时问题变得非常复杂。如果你在选择答案时完全不考虑每个答案对应的概率，仅仅随机地猜测答案，则问题还原为已经讨论过的情况；如果你把不同的概率考虑进去，情况就不同了。假设有5个不确定的答案，其中3个是正确的，2个是不正确的，而每个答案正确的概率不同：A是正确答案的概率为90%，B是正确答案的概率为75%，等等。由于在5个答案中有3个答案是正确的，所以这5个对应的概率值的总和为300%。如果你猜测A是正确答案，并且事

实证明你的猜测是正确的，则在剩下的4个答案中还有2个正确答案和2个不正确答案，于是剩下的4个答案对应的概率值的总和变为200%。请注意，B对应的概率已不再是75%。事实上，剩下的4个答案对应的概率都已受到影响。对于这个问题的分析极其复杂，我们不作进一步的讨论。总之，在这一类问题中我们假设以下四个条件是成立的：

（1）两个参与者的知识范围完全相同；

（2）对于每个答案，参与者要么确知其正确与否，要么完全随机地猜测；

（3）每个参与者知道 x 个正确答案和 y 个不正确答案；

（4）每个参与者都选择最佳策略。

约定这四个条件是为了计算的方便，实际情况要比约定的情况复杂。下面是针对不同的 x和 y 的值的讨论。计算中做了一些近似处理，但精度是足够的。

在这类问题中，确切地知道某个答案不是正确答案也是有用处的。如果你确切地知道某4个答案不是正确答案，也就等于你确切地知道了哪8个答案是正确答案。所以，当 $x=8$ 或 $y=4$ 时，问题很简单，你应当投标8，并且可以保证100%获胜。当 $x=7$ 时，无论 y 是多少你都应当投标8。当 $x=6$ 时，你总是应当投标7。此时对手的最佳反应取决于 y 的值，如果 $y=2$，则对手应当投标8；否则对手不应当加码。当 $x=5$ 时，你的行动取决于 y 的值：当 y 较小时你应当投标6，而当 y 较大时你应当投标7。此时对手通常不应加码。当 $x=4$ 时，最后的投标额肯定是6。当 y 较小时，最佳的游戏进程是你投标5而对手加码到6；当 y 较大时，最佳的游戏进程是你投标6而对手不再加码。当 $x=3$ 并且 $y=3$ 时，你应当投标6；当 $x=3$ 而 y 较小时你的投标额应当小一点。当 $x=2$ 时，你应当投标4

或5。当 $x=1$ 时，如果 y 较小就投标3，如果 $y=3$ 则可以投标5。最后，当 $x=0$ 时，虚张声势的心理战术对你很合适。此时你的投标额应当比 $x=1$ 时的投标额小1。

请注意：在实战中不要拘泥于以上结论。以上结论依赖于若干特定的条件，在实战中这些条件通常是不具备的，所以在实际情况中的最佳策略可能与理论计算的结论不同。然而，以上结论在实战中可以有非常实用的指导价值。如果你确知8个正确答案或4个不正确答案，你当然应当投标8。在其他情况下，你的投标额应当比你确知的正确答案数大1或2，而在你确知的不正确的答案比较多时你应当更加自信。如果你第一次投标的投标额较低，你就把主动权交给了对手，所以正确的做法总是"一次叫足"。假如你希望投标6，那么第一次你就应当投标6。如果你先是投标4，并且准备在对手投标5时加码到6，这种策略肯定是不对的，因为对手可以直接投标6，你就处于被动地位。在对手投标之后，如果你觉得你加码之后成功的概率高于50%，你就应当毫不犹豫地加码。当对手成功的概率较高时（比如说超过60%），加码通常是正确的，即使加码以后你的胜算不足50%。最后需要注意的是，心理战术是很有用的。精通概率并不足以取胜，据我所知，最优秀的数学家在赌场上通常无法对抗资深赌徒。

下面我们研究游戏的第一阶段。这个阶段的目的是淘汰3名参与者中的1个。游戏的形式是主持人提供16个对象，其中有11个具备相同的性质。参与者的任务是挑出具有这个性质的对象。第一个参与者先回答，他每答对一个答案，他的奖金都会增加。在每一次正确回答之后，他有权选择继续回答或把回答问题的权利交给下一个参与者。一旦他的回答错误，他已取得的奖金就被清0，由下一个参与者继续回答。如果所有

168 正确答案都已被挑出，或者所有不正确的答案都已被标明，则游戏结束。最终，得到奖金最少的参与者将被淘汰。如果有两个人的奖金并列最低，则淘汰回答出正确答案较少的一个。

在游戏结束时，有两个（甚至全部三个）参与者的奖金总额为0的情况司空见惯，因为“清0”事件发生的频率非常高。聪明的战略是在游戏过程中争取尽量多地回答问题，不必在乎清0的风险。在游戏的初期更应当积极地回答问题。这种战略的根据在于，当游戏结束时你与另一个人的奖金并列最低的概率非常高，回答较多的问题有助于逃脱被淘汰的命运。你的主要目标是进入下一阶段的角逐，而不是在这一阶段赢得零星的奖金。与进入下一阶段游戏的利益相比，在这一阶段中的得失微不足道。

下面描述一个典型的游戏历程。主持人给出16个候选答案，其中有11个是汉诺斯菲尔德东南部村庄的名称，另外5个不是。16个候选答案如下，其中的错误答案被标上了星号：①

A	B	C	D
E*	F	G*	H
I	J	K*	L
M*	N*	O	P

甲：（我对这几个地名一无所知，碰碰运气吧。）

① 原文中给出的是英文地名，我们对具体的地名不感兴趣，为了方便阅读仅以英文字母代替——译者注。

"H。"——"回答正确，得 10 英镑。"

(甲继续回答。)"I。"——"正确，再加 20 英镑，总共 30 英镑。"

(甲继续回答。)"N。"——"很遗憾，你被清 0 了。"

乙："A。"——"回答正确，得 30 英镑。"

(乙继续回答。)"D。"——"正确，再加 40 英镑，总共 70 英镑。"

"我放弃。"——"乙得 70 英镑。下面轮到丙回答。"

丙："M。"——"回答错误。下面轮到甲回答。"

甲：(为什么乙要放弃呢?)"B。"——"回答正确，得 50 英镑。"

"P。"——"正确，再加 60 英镑，总共 110 英镑。"

"O。"——"正确，再加 70 英镑，总共 180 英镑。"

游戏按这个方式进行。假设若干轮之后出现此局面：甲答对的问题多于乙，乙答对的问题多于丙；甲、乙、丙已获得的奖金分别为 40、0 和 70 英镑；游戏已进行到最后一轮，在这一轮中将有 1 人被淘汰；已有 7 个正确答案和 2 个错误答案被标明；每个参与者随机地在候选答案中选择，关于这些地名的实际知识不起作用。在这种情况下，真正需要技巧的仅仅是这个问题：当你做出一次正确的回答之后，你应当继续回答还是弃权？现在丙刚刚弃权，又轮到甲回答。 169

甲：(还剩下 7 个候选答案，其中 4 个正确，3 个不正确。概率对我有利。)"E。"——"回答错误，你被清 0 了。"

乙：(我回答正确的概率是 4/6。)"G。"——"回答错误，你被清 0 了。"

丙：(我的机会高达 4/5。)"L。""回答正确，加 80 英镑。"

现在甲和乙的奖金额都是 0，而丙的奖金额为 150 英镑，还剩下 3 个正确答案和 1 个错误答案。3 个正确答案的价值分别是 90、100 和 110 英镑。你认为游戏应如何进行？每个参与者被淘汰的概率分别是多少？此时丙是否应继续回答？

如果丙相信甲的智力，他应当放弃回答。理由如下：此时甲和乙奖金相同（都是 0），而甲回答出的正确问题多于乙，一旦甲选中了 1 个错误答案，则所有错误答案都已被标明，游戏结束，乙被淘汰（因为乙和甲的奖金并列最低而乙回答出的正确问题少于甲）。所以，甲非常乐于挑出 1 个错误答案。甲的最佳策略是一直回答下去，直到游戏在他的手中结束。如果甲的第一次回答是正确答案，甲不能放弃。一旦甲在这个时刻放弃，他就有可能被淘汰——如果此后乙和丙都避免选中错误答案，这种情况就会发生。不过，如果甲连续选中 2 个正确答案，此时放弃是安全的，因为甲已经得到了 90 + 100 = 190 英镑，乙无论如何努力都无法超越甲和丙的成绩，乙被淘汰的命运已无法避免。结论：如果每个人都追求自己的最大利益，则丙应放弃，甲应坚持回答，乙最终将出局。

如果丙放弃，但甲犯了错误——他在答出一个正确答案之后放弃，结果如何呢？在这种情况下，每个参与者都有可能被淘汰，而且每个人被淘汰的概率相等。（此时剩下 3 个答案，2 个正确的和 1 个错误的，选中错误答案者即遭淘汰。）

如果丙犯了错误——他没有放弃，而是继续回答问题，结果如何呢？丙有 1/4 的可能被淘汰。在这个终局，甲始终居于主动地位，因为甲回答出的正确答案较多。此时丙可能会懊悔。如果丙在游戏早期积极地回答问题，以确保自己回答出的正确答案较多，终局形势就非常简单而有利——他只要持续不断地回答问题直到游戏在他手中结束，就可以保证顺利过关。

这个例子的目的在于说明为什么在游戏的早期应当积极地回答问题。更多地回答问题可以保证回答出的正确答案较多，这样在终局就可以保持主动。不必在意辛辛苦苦积累起来的奖金一下子被清0。事实是：无论你已经积累了多少奖金，都将被一次错误回答清0，这是游戏规则规定的，你无法通过谨慎和智谋避免。而在接近终局时，你应当小心地计算各种策略对应的成功的概率，就像在上例中甲所作的分析。在上例中，甲其实根本不关心自己的回答是否正确，他只要坚持回答下去就可以把乙淘汰。至关重要的因素是此时乙的奖金额为0。如果乙的奖金额不是0，情况则大不相同——哪怕乙的奖金额仅为10英镑。接近终局时冷静的分析和缜密的推理是很重要的。 170

在游戏的初期，可能出现很多种情况，这使得终局局面千差万别。不同的终局局面对应着不同的最佳策略，所以很难得出一个处理这个问题的万能公式。不过我们可以找到对付这个问题的一般方法，其核心为“倒推法”。为了说明这种方法，我们首先研究简化后的情况。假设游戏规则不变，但参与者只有两个：甲和乙。游戏的目标是把对手淘汰，而不在乎在游戏中具体赢得多少奖金。最后一个问题的价值是100英镑，倒数第二个问题的价值是90英镑，每个问题的价值比后一个问题少10英镑。

我们的目标是发现在各种局面下甲胜出的概率以及甲的最佳策略。每次回答都依靠随机猜测，参与者可以选择的行为是“放弃”或“继续”。我们用（a，b，c，d）表示游戏达到的局面，其中：

a表示剩下的正确答案数；

b表示剩下的错误答案数；

c表示甲当前的奖金额；

d表示乙当前的奖金额。

例如，(3，2，70，60) 表示还剩下5个候选答案，其中3个正确，2个错误，下一个正确答案的价值为80英镑，甲当前奖金额为70英镑，乙当前奖金额为60英镑。甲面临“放弃”或“继续”的选择。甲应如何选择？他胜出的概率是多少？

游戏结束的条件是所有正确答案都已被挑出或所有错误答案都已被标明，即a和b中有一个为0。(2，0，120，0) 表明甲获胜，(0，1，90，100) 表明乙获胜。如果游戏结束时出现两人奖金额相同的情况，例如 (2，0，0，0)，则胜负取决于谁回答出的正确问题较多。在终局局面中，a和b中有一个为0。对于每个终局局面，我们可以标明甲获胜的概率。当c大于d时，标明概率为1；当c小于d时，标明概率为0；c等于d时，胜负取决于谁回答出的正确答案较多，此时标明概率为p。这样，每一个终局局面都可以标明1、0或p。我们的目标是对于a和b不为0的情况也标明这个概率值。前提假设是双方的目标都是增加自己获胜的概率。

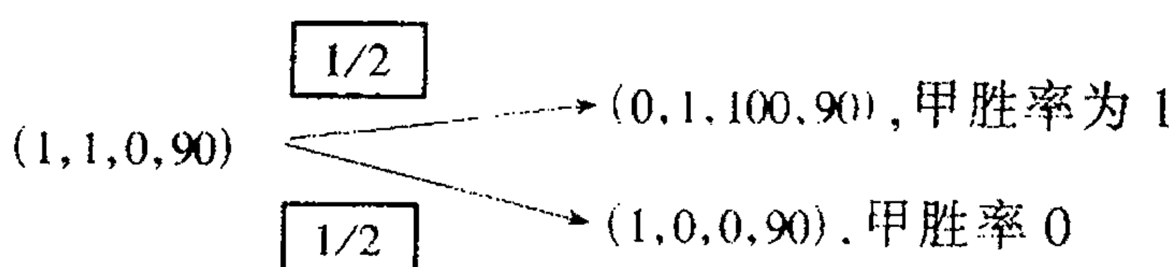
从局面 (3，2，70，60) 开始讨论。假如甲决定继续，但选中1个错误答案。这种情况发生的概率为2/5，局面转变为 (3，1，0，60)，乙开始回答问题。乙回答正确的概率为3/4，此时局面转变为 (2，1，0，140)，乙可以选择放弃或继续。假如乙决定继续，他有1/3的概率猜错，局面转变为 (2，0，0，0)，游戏结束。整个过程表示如下：

$$(3,2,70,60)\xrightarrow{2/5}(3,1,0,60)\xrightarrow{3/4}(2,1,0,140)\xrightarrow{1/3}(2,0,0,0)$$

为了求出在某一特定局面下的最佳策略，我们需要列出这一局面可能演变出的各种结果。由于游戏结束时双方获胜的概

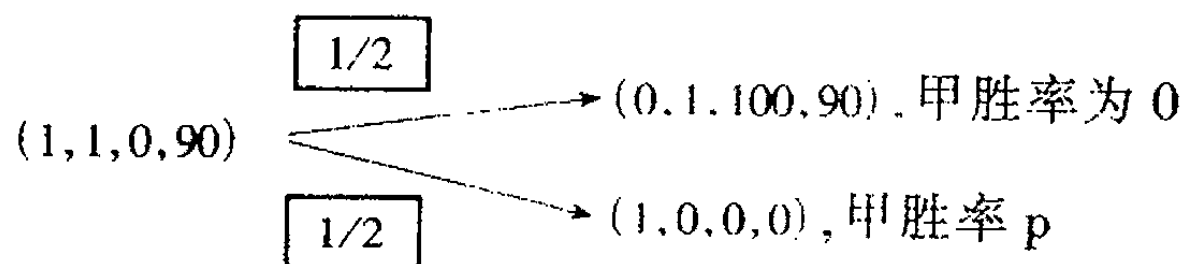
率是已知的，而每种情况发生的概率可以计算，所以我们可以利用“倒推法”从后向前计算出各种局面下甲的胜率。第一步是研究只剩下两个候选答案的情况，游戏的进行有两种可能结果：回答正确或回答错误。当甲（或乙）回答问题时，有 1/2 的概率赢得 100 英镑的奖金，同时有 1/2 的概率被清 0，无论哪种情况发生，甲获胜的概率是可以确定的。由于两种结果出现的概率相同，所以甲的胜率就是这两种结果对应的胜率的平均值。当一个参与者面临“放弃”或“继续”的抉择时，他应当计算出每一种抉择对应的胜率，而选择胜率较高的一个。

例如，局面为（1，1，0，90）。如果轮到甲回答问题，游戏的进程如下：



显然，甲获胜的概率为 1/2。

如果轮到乙回答问题，游戏的进程如下：



此时甲的胜率为 p/2。如果此时乙可以选择“放弃”或“继续”，他的决定应当取决于 1/2 和 p/2 的比较结果，他应当选择值较小的一个（因为这个值表示甲的胜率）。

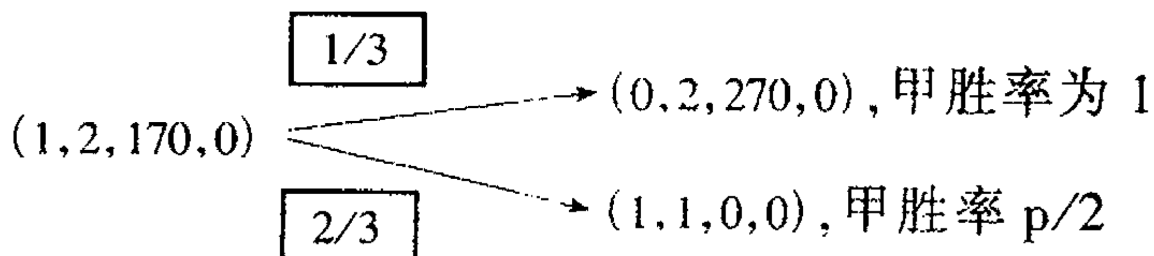
利用这种方法可以处理所有剩下的答案只有两个时的局面。下面我们考虑剩下 3 个答案，其中 1 个正确而 2 个错误时的局面。此时，如果一名参与者选中了正确答案（概率为 1/

3)，则游戏结束；如果一名参与者选中了错误答案（概率为2/3)，则游戏转变为只剩下两个答案的情况。无论哪种结果发生，甲获胜的概率都是可以决定的。利用类似的方法可以算出这一局面下甲的胜率。当一名参与者面临抉择时，他应当计算每一种抉择对应的胜率，选择对自己有利的一种情况。

下面是一个例子：

局面为（1，2，170，0)．甲可以选择“放弃”或“继续”。甲如何选择？

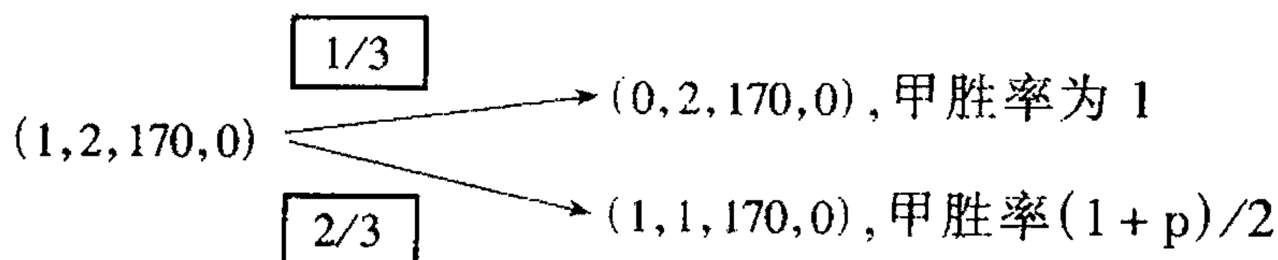
如果甲选择继续，游戏进程如下：



因此，如果甲决定继续，其胜率为

(1/3) × (1) + (2/3) × (p/2) = (1+p) /3。

如果甲选择继续，游戏进程如下：



173

因此，如果甲决定放弃，其胜率为

(1/3) × (1) + (2/3) × (1+p) /2= (2+p) /3。

无论 p 为什么值，第二个值总是比第一个值大，所以在此局面下甲的最佳策略是放弃。

利用同样的方法可以计算剩下 2 个正确答案和 1 个错误答案的局面，并为具体的局面标明甲的胜率。此后，我们可以把问题推广到剩下 4 个答案以及剩下 5 个答案的局面。我们最初

考虑的局面是（3，2，70，60），即剩下 3 个正确答案和 2 个错误答案的情况。一步一步地应用“倒推法”，计算思路很简单，但计算过程很繁琐。这种任务最好交给计算机完成。表 9.1 是计算结果。如果双方都按最佳策略行事，p 的值与双方的奖金额无关，始终等于 40%。

在表 9.1 的前两种情况下，甲没有选择“放弃”或“继续”的权利，他必须回答问题。（只有在完成一次正确回答以后才有选择的权利，这意味着当前的奖金额不能为 0。）在第三种情况下，放弃和继续的结果是一样的。在最后两种情况下，甲应当选择放弃。我要再次强调，这些结论依赖于一个基本假定：每个参与者完全随机地在候选答案中挑选答案。如果这个条件不被满足，结论会不同。在游戏结束时，双方的奖金额经常相同，此时胜负取决于谁回答出的正确答案较多。因此，在游戏早期应当积极地回答问题，以保证终局时的主动。

表 9.1：剩下 3 个正确答案和 2 个错误答案时甲的胜率

甲的当前奖金额	乙的当前奖金额	甲回答出的正确答案较多时的结果	乙回答出的正确答案较多时的结果
0	500	60%	20%
0	0	70%	30%
70	60	70%	30%
70	500	70%	30%
130	0	80%	40%

174

当然，这些技巧在实战中通常是没用的。在实战中，参与者没有时间精确地计算各种概率，必须干脆利落地做出决定。以上分析的目标只是向读者展示概率的知识如何帮助我们处理

类似的问题。比如说，10 年以后你将退休，现在你需要决定是否开始购买养老保险。你可以利用“倒推法”做决定。首先你可以计算出在第 10 年买保险的得失，然后根据第 10 年的计算结果决定第 9 年的得失，下一步又可以根据第 9 年的计算结果决定第 8 年的得失，最后你就可以决定现在的最佳选择。

另一个问题与此类似。你要踩着几排石头过河，每块石头与对岸的距离不等，石头在水面上的高度也不同。你应当如何决定前进的路线？当你距离对岸很近时，你可以很容易地决定路线。下一步，你可以根据刚刚得到的结论决定距离对岸稍远时的路线。不断应用“倒推法”就可以决定第一步应当如何走。

幸运转盘

这个游戏是“奖金揭晓”节目的高潮。游戏由三个人——甲、乙、丙——参与，最终获胜者可以得到一笔奖金。游戏通过旋转一个大转盘进行，转盘上标有数字 5，10，15，20，……，100，每次旋转可以随机地得到其中一个数字。甲先旋转转盘，如果他对得到的数字满意，则这个数字就是他的得分，接下来由乙旋转转盘。如果甲对这个数字不满意，他可以要求再转一次，他的得分就是两次得到的数字的总和，此后把转盘交给乙。乙和丙以同样的方式参与游戏，每个人的得分是公开的。游戏的目标是使自己的得分最接近于 100，但不能超过 100。得分超过 100 的人自动淘汰。如果甲和乙都遭淘汰，则丙自动成为胜利者。如果出现两个人得分相同的情况，由加赛分出胜负。加赛的方式是两人分别旋转转盘，谁得到的数字

大谁就获胜。

如果你参与游戏，你将选择旋转一次还是两次？当你的得分与前一个人相同时，你是否应当再转一次？出场顺序对结果是否有影响？谁获胜的概率最大？ 175

这个问题与“猜猜看”类似，但计算方法更简单。首先我们考虑简化后的问题。假设游戏规则不变，而参与者只有甲和乙两人。为了确定甲的最佳策略，我们先要研究乙对于各种情况的反应。

如果甲的分数超过 100，游戏自动结束，乙获胜。在其他情况下，甲的得分在 5 和 100 之间，乙出场。通常乙的策略是明确的：如果乙第一次得到的数字比甲的得分小，他必须转第二次；如果乙第一次得到的数字比甲的得分大，他不必转第二次就已经获胜。如果乙第一次得到的数字和甲的得分相同，他应如何选择呢？他将这样考虑：“如果我不转第二次，胜负通过加赛的结果决定，每个人的机会都是 1/2。所以，如果转第二次能使我的胜率超过 50%，我就应当继续转。转第二次的结果是 20 个数字中的 1 个。如果我的当前得分不超过 45，则转第二次使得我的分数超过 100 的概率低于 50%，此时我应当继续转；相反，如果我的当前得分不低于 55，则转第二次使得我的分数超过 100 的概率高于 50%，此时我应当停止；如果我的当前得分为 50，则无论怎样选择结果相同。所以，当我的当前得分不超过 50 时，我决定再转一次，否则我愿意通过加赛决胜。”

下面考虑甲的策略。甲完全可以推测出乙的策略。甲有两个选择：转一次或转两次。如果甲转一次，则他第一次得到的分数越高，获胜的概率就越大；如果甲转两次，他必须考虑两次得到的数字之和超过 100 的危险，他第一次得到的分数越

高，这种危险就越大。随着甲第一次得到的分数的增加，有两个相反的因素起作用，当第一次得到的分数达到一定值时，这两个因素取得均衡。这个值是可以计算的，事实上，它就是50。于是，当甲第一次的得分不超过50时，他应当转第二次；而当甲第一次的得分超过50时，他就应当停止。例如，当甲第一次的得分为55时，如果甲停止，他的胜率为29%，但如果他转第二次，他的胜率就下降为28%。总的说来，采用这种策略可以保证甲的胜率达到46%。也就是说，甲有54%的概率失败。

下面我们考虑参与者为三人的情况。我们还是从最后一个参与者（丙）的立场开始讨论。丙的直接对手是甲和乙中得分较高的一个。惟一需要仔细研究的即丙第一次的得分与某个人并列第一时的情况。此时需要考虑两种可能：第一，甲和乙中只有一个人的得分与丙第一次的得分相同；第二，甲和乙的得分相同，并且与丙第一次的得分相同。在第一种可能下，如果丙停止，则加赛将在两个人之间进行，情况还原为两个人参与游戏的情况，此时丙的最佳策略不变：当得分不低于50时继续转第二次，否则停止。在第二种可能下，如果丙停止，则加赛将在三个人之间进行，丙会这样考虑："如果胜负由加赛决定，则三个人胜出的概率相等，即我只有1/3的机会。所以，只要我转第二次的胜率不低于1/3，我就可以决定转第二次。如果当前得分为65，则转第二次的胜率为7/20（在5到35之间有7个数字），刚好大于1/3；如果当前得分为70，则转第二次的胜率低于1/3。所以，我的决定是：当第一次的得分不高于65时，转第二次；否则停止，在加赛中争取机会。"

下一步考虑乙的策略。乙完全可以推出丙的策略。如果乙第一次得到的分数低于甲的分数，毫无疑问乙应当转第二次；

如果乙第一次得到的分数高于甲的分数，甲已经被淘汰，问题转化为只有两人参与的情况，乙的策略我们已经确定了（即当乙第一次的得分不超过50时，转第二次，否则停止）；如果乙第一次得到的分数恰好等于甲的分数，乙应如何抉择？

乙的当前得分越高，转第二次使得分超过100的风险就越大。与刚才我们讨论过的情况一样，乙抉择的依据是如何使自己的胜率不低于通过加赛取胜的胜率。结论是：如果当前得分不高于65时，转第二次；否则停止。

最后一步考虑甲的策略。甲完全可以推出乙和丙的策略，当然，计算工作应当事先完成。计算过程相当繁琐，但结论很简单：当甲第一次的得分不超过65时，他应当转第二次；而当甲第一次的得分超过65时，他就应当停止。当甲第一次的得分为70时，如果甲停止，他的胜率为21.5%；而如果他转第二次，则他的胜率下降到17%。

现在我们已经知道了三个人的最佳策略。丙在游戏中只有优势，但优势很微弱。如果三个人都按最佳策略行事，则甲的胜率略低于31%，乙的胜率接近33%，丙的胜率略高于36%。

实际的计算过程相当繁琐，我略过了大部分细节。然而计算方法并不复杂，主要原则是列举出各种情况的概率。以上讨论假设转盘给出每个数字的概率相等，如果这个条件不被满足，我们的结论可能构成严重的误导。

177

抢答游戏

抢答游戏的规则比较复杂。参与者通常为16～18岁的孩子。游戏需要一块计分板，如图9.1所示。计分板由若干个六

边形组成，纵向m个，横向n个。图9.1中m=4，n=5，不过实际比赛中m和n可以取其他值。每个六边形对应一个问题，问题将由参与者抢答。每个六边形上标有一个字母，这个字母既是此六边形的标记，同时也是答案的第一个字母。字母的排列方式是在游戏开始之前任意安排的。

下面是游戏的一个片断：

主持人："恭喜你，理查德，回答正确。请你指定下一道题。"

理查德："p。"

主持人："p对应的问题是：有一种符号为K的化学元素以p开头，这种化学元素是什么？"

（抢答器的声音响起。）主持人："维拉斯先抢答。请维拉斯回答。"

维拉斯："钾。"

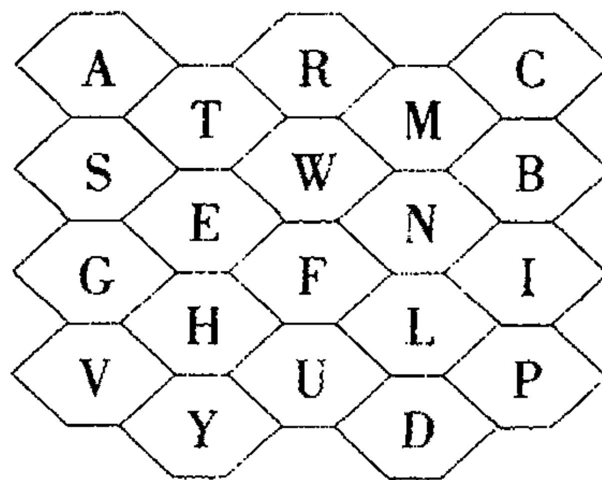
主持人："回答正确。现在我们把p涂成红色。请维拉斯指定下一道题。"[①]

在游戏开始时，计分板上的六边形没有着色，仅标有字母。而一旦一个六边形对应的问题被正确回答，这个六边形就被涂上蓝色或红色。三个参与者分为两队，蓝队两人，红队一人。蓝色代表蓝队答对的问题，红色代表红队答对的问题。蓝队的目标是使蓝色的六边形横贯计分板，红队的目标是使红色的六边形纵贯计分板，不考虑各个问题出现的顺序。一旦某个队完成目标，游戏结束。第一个六边形是随机选定的，其后每

① 钾的西文名称为Potassium，元素符号为K——译者注

个六边形由刚刚做出正确回答的参与者指定。如果被选中的问题三个人都不会，主持人会提供一个备用的问题，备用的问题指向同一个六边形。

图 9.1



178

每答对一个问题可以获得 5 英镑的奖金，不过我们忽略这点微不足道的奖金，目标只有一个：取得本局的胜利。游戏采取三局两胜制，先胜两局的队伍进入下一轮，另一方则被淘汰。为了尽量缩短游戏时间，有一条补充规定：如果选择某一个六边形使得某队有可能立即取胜，则必须选择此六边形。

每局游戏一定可以分出胜负。必然有一个队完成目标，而且仅有一个队可以完成目标。这个问题很明显，读者可以自行证明。

这个游戏对双方公平吗？蓝队有两个人，而红队只有一个人，所以蓝队抢答成功的概率是红队的二倍。不过作为补偿，蓝队完成目标的难度也比较大，因为横贯计分板至少需要答对 5 个问题，有时甚至需要答对 7 个，而红队答对 4 个问题就有可能获胜。假定每个参与者的知识水平相同，抢答成功即意味着回答正确，于是对于每个问题蓝队抢答成功并回答正确的概率为 2/3，而红队抢答成功并回答正确的概率为 1/3。在此前

提下，每局中双方的胜率各为多少？在三局两胜制中，双方赢得整场胜利的概率各为多少？

在实际的游戏中，一旦某方完成目标则游戏结束。不过为了便于研究，我们假定每一局都抢答完全部的 20 个问题。这一假定不会改变游戏的结果，原因我已经说过：最终只有一个队能够完成目标。如果一个队已经完成目标，则对方在计分板上的通路已被切断，完成目标已不可能。这样假定的好处是每一局中抢答的次数相同，便于用数学手段处理。这与《三局两胜》一章中的情况相同。

179

图 9.2

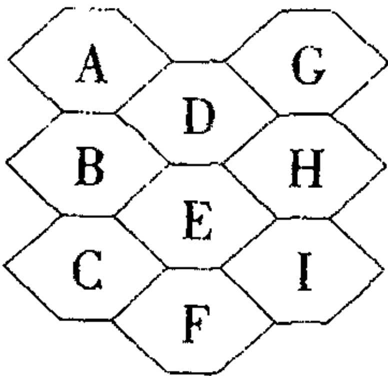


图 9.1 表示的计分板比较复杂，直接入手比较困难。所以我们先研究一个简化后的情况。假定 m 和 n 都等于 3，计分板如图 9.2 所示。

这个游戏显然对蓝方有利。如果蓝方取胜，则蓝方至少答对 3 个问题。有 9 种 3 个问题的组合可以使蓝方获胜，即

ADG　ADH　BDG　BDH　BEH　BEI　CEH　CEI　CFI

同样，如果红方取胜，则红方至少答对 3 个问题。但是只有 7 种 3 个问题的组合可以使红方获胜，即

ABC　DBC　DEC　DEF　DEI　DHI　GHI

这意味着，蓝队在 9 个问题中仅答对 3 个但最终获胜的概率高

于红队在同样情况下的获胜概率。同样，蓝队在9个问题中仅答对4个但最终获胜的概率高于红队在同样情况下的获胜概率——有46种4个问题的组合可以使蓝队获胜，而只有39种4个问题的组合可以使红队获胜。类似地，有85种5个问题的组合可以使蓝队获胜，而只有80种5个问题的组合可以使红队获胜；有77种6个问题的组合可以使蓝队获胜，而只有75种6个问题的组合可以使红队获胜。所以，计分板的形状对蓝队有利。在图9.1中，也有同样的现象。

假设蓝队和红队都由1名队员构成，采用图9.1为计分板，双方队员的知识水平和反应速度相同，则双方的胜率如何？根据假定，每个六边形被涂成蓝色或红色的概率相同。每 180 个六边形的最终颜色有两种可能，总共有9个六边形，所以全部的可能性为512种（2的9次方），每种可能性出现的概率相等。其中247种使得红色六边形纵贯计分板，而265种使得蓝色六边形横贯计分板。因此，蓝队的胜率为265/512，约等于52%。如果计分板的形状改为55（即 m 和 n 都等于5），则蓝队的胜率增加到53.5%。

现在回到最初的问题。红队赢得每个问题的概率只有1/3。我们首先计算一种典型的状态：红队恰好赢得9个问题而蓝队恰好赢得11个问题。此时计分板上可能的颜色分布有很多种，其中的某些颜色分布使红队获胜，其他的颜色分布使蓝队获胜，两个队的胜率的比例就是两种颜色分布的比例。计算过程非常复杂，因为需要考虑的因素太多了。但是不必担心，我们可以请计算机帮忙，计算机处理这个问题很轻松。计算的结果是：一共有167 960种可能的颜色分布，其中的70 964种使红队获胜，而96 996种使蓝队获胜。由于每个六边形被涂以红色的概率为1/3，而被涂以蓝色的概率为2/3，每种颜色分布

出现的概率为

$$(1/3)^9 \times (2/3)^{11}$$

于是，红队通过赢得 9 个问题取胜的概率为

$$70\ 964 \times (1/3)^9 \times (2/3)^{11} = 0.0417$$

当红队赢得的问题在 4 和 15 之间时，红队的胜率可以用类似的方法求出；当红队赢得的问题低于 4 时，红队的胜率为 0；当红队赢得的问题高于 15 时，红队的胜率为 1。把这些概率加起来，就得到红队在这个游戏中的胜率：0.1827。另外，为了赢得整场游戏的胜利，蓝队必须在三局中拿下两局，所以红队最终胜出的概率比这个值还要低。如果不经过仔细地计算，你可能以为纵向的六边形较少可以抵偿红队人数较少的缺陷，但实际情况是红队的机会小得可怜。

关于图 9.2 的分析表明，横向贯通计分板比纵向贯通计分板容易。由于蓝队在人数上已经占有优势，所以如果把计分板改为横向窄而纵向长，并把蓝队的目标改为纵向贯通，红队的劣势会小一点。例如，当 $m = 5$ 而 $n = 4$ 时，红队的胜率为 0.224 5；当 $m = 5$ 而 $n = 3$ 时，红队的胜率为 0.322 2。

181

在上一章我们研究过如何通过每局的胜率计算整场的胜率。当每局的胜率接近 20% 时，在三局两胜制下整场的胜率接近 10%。利用我们研究过的知识可以算出各种局面下双方的胜率。以上讨论回避了某些细节，比如每个参与者在反应速度、知识水平等方面的差异，这些因素对游戏结果可以产生重要的影响。

在计分板上，有一些六边形比其他的六边形更重要，如果占领了这些六边形，获胜的概率会增加。例如，图 9.1 中的 F 就是一个重要的位置，而 Y 和 D 的作用就比较小。但非常有趣的是某一方赢得某个位置的概率与这个位置的重要性并无关

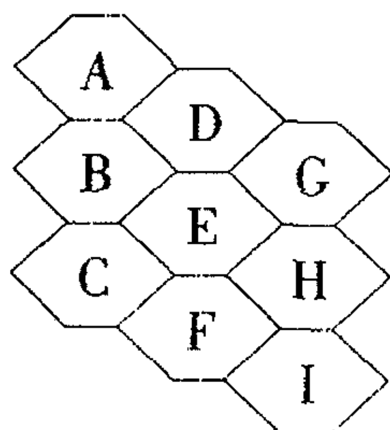
系，而且赢得一个位置总是没有任何坏处的。在游戏中，每个参与者都有机会选择下一个六边形的位置。从赢得本局游戏的立场出发，你可以胡乱选择，因为如何选择对胜负没有影响。但是我建议你考虑另一个因素：每答对一个问题可以得到5英镑的奖金。所以无论对于哪一方来说，过早结束游戏都不是明智的选择，每减少一个问题就意味着参与者的奖金收入减少5英镑。双方在选择下一个位置时都应当考虑如何可以使游戏的进程尽量地延长。为了达到这个目的，应当尽量选择那些不重要的位置，比如图9.1中的Y、D、R和A。

附 言

有一种游戏与抢答游戏类似，即六边形游戏。派特·汉恩（于1942年）和约翰·纳什（于1948年）各自独立地发明了六边形游戏。六边形游戏的棋盘如图9.3。棋盘由六边形构成，纵向m个，横向n个，图中表示的是m和n都为3的情况。请比较与图9.2的差别。

182

图9.3



在六边形游戏中，双方（蓝方和红方）轮流在棋盘上涂

色，每次只能涂一个六边形，可涂以蓝色或红色。蓝方的目标是使蓝色横贯棋盘，红方的目标是使红色纵贯棋盘。这个游戏与抢答游戏的差别在于，在这个游戏中横向和纵向获胜的机会是相等的。例如，一共有 8 种 3 个六边形的组合可以使蓝方获胜：

ADG BDG BEG BEH CEG CEH CFH CFI

同样，一共有 8 种 3 个六边形的组合可以使红方获胜：

ABC DBC DEC DEF GEC GEF GHF GHI

在六边形的组合的个数不为 3 时情况相同，并且对于任意的 m 和 n 的值，只要 $m = n$，这个结论也成立。无论 m 和 n 取什么值，这个游戏一定可以分出胜负，有且仅有一方可以完成目标。数学家已经证明，当 $m = n$ 时，在这个游戏中先行的一方可以保证必胜。然而，当 m 和 n 的值很大时，发现必胜的策略非常困难。

把抢答游戏改造成一个公平的游戏很简单。只要把计分板改为六边形游戏的形状，并使得 $m = n$、双方的队员人数相等就可以了。不过为了便于电视转播，我们可能需要把电视机的屏幕改成菱形。

第十章　赌场中的游戏

1960年英国政府颁布了《赌博与竞技法案》。法案的目的是规范赌博业，但同时带来了赌场的兴起。1968年的《竞技法案》增加了对赌场的限制，包括对赌场的营业地点的规定。赌场中允许开展什么样的游戏有明确的规定，地方官要对赌场进行严格审查，甚至要对赌场的从业人员的品行进行考核。目前，每年在赌博业流通的资金大约为26亿英镑，其中有4.7亿英镑在赌场内流通。伦敦有20多家赌场，占全国的营业额的80%。

立法机关对赌博业的态度始终不明朗：他们允许赌博业存在，但不允许赌博业明目张胆地存在。1960年，立法机关意识到禁止赌博是不可能的，于是出台了相关法案。与其禁止赌博，不如进行规范，并且向赌博业征税。英国的赌场始终受到严格的限制。最初赌场不允许进行广告和促销活动，但大英帝国彩票兴起之后，赌场老板乘机摆脱了一些限制。现在赌场可以做广告了，但对广告的内容依然有限制。在广告中可以提及赌场有多少赌台，但不能提及赌注的金额。总之，不能公开吸引大众参与赌博。赌场可以向顾客提供饮食服务，但不允许兼

营娱乐服务。与拉斯维加斯的同行相比，英国的赌场相当地单调和保守。

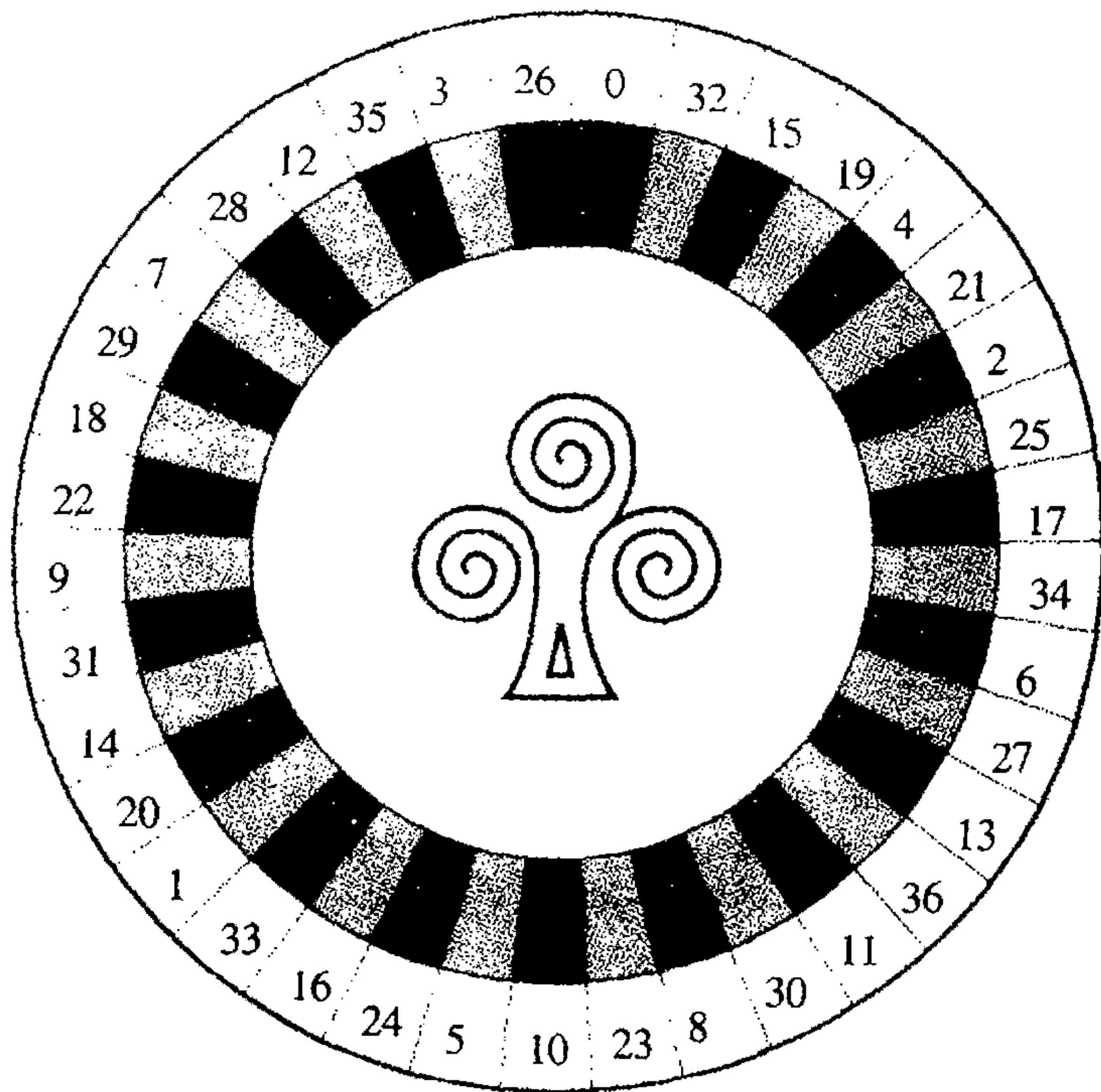
各种游戏都被设计成对赌场有利的形式，这是赌场重要的盈利手段。不过对每一个顾客来说，这些不利因素非常微弱，可以忽略。而且，顾客可以在其他方面获得补偿，比如结识朋友，消磨时间等等。本章将讨论一些在赌场中经常遇到的问题，比如在某个游戏中的收益率是多少，是否有提高获胜概率的技巧，在某个局面下是否应当加注，等等。全部讨论基于英国赌场的规则。

本章的目的不仅是介绍一些实战技巧，更重要的是展示概率知识如何应用于赌博活动。在多数情况下，计算的过程相当复杂。你当然可以略过乏味的计算，但是我建议你仔细推敲全部你能读懂的内容。

轮盘赌

到目前为止，轮盘赌是最流行的赌场游戏，吸引了全部赌注的 60%。标准的英国轮盘上标有 37 个数字：从 0 到 36。数字 0 被涂成绿色，其他数字则红黑相间，如图 10.1 所示。玩家在进入游戏以前要买一些筹码，在同一个赌台上不同玩家的筹码颜色不同，这样可以避免混乱和争议。在同一家赌场中，有些赌台要求的最低下注额非常低，另外一些赌台则要求较高。关于最高下注额也有规定，通常是最低下注额的 100 倍。

图 10.2 是下注区，上面标明了赌场的赔率。图中的 12 个字母代表 12 种可以下注的位置。A 代表在单个数字上下注，赔率为 1 赔 35；B 代表在两个数字上下注，赔率为 1 赔 17；C



186

图 10.1:标准轮盘

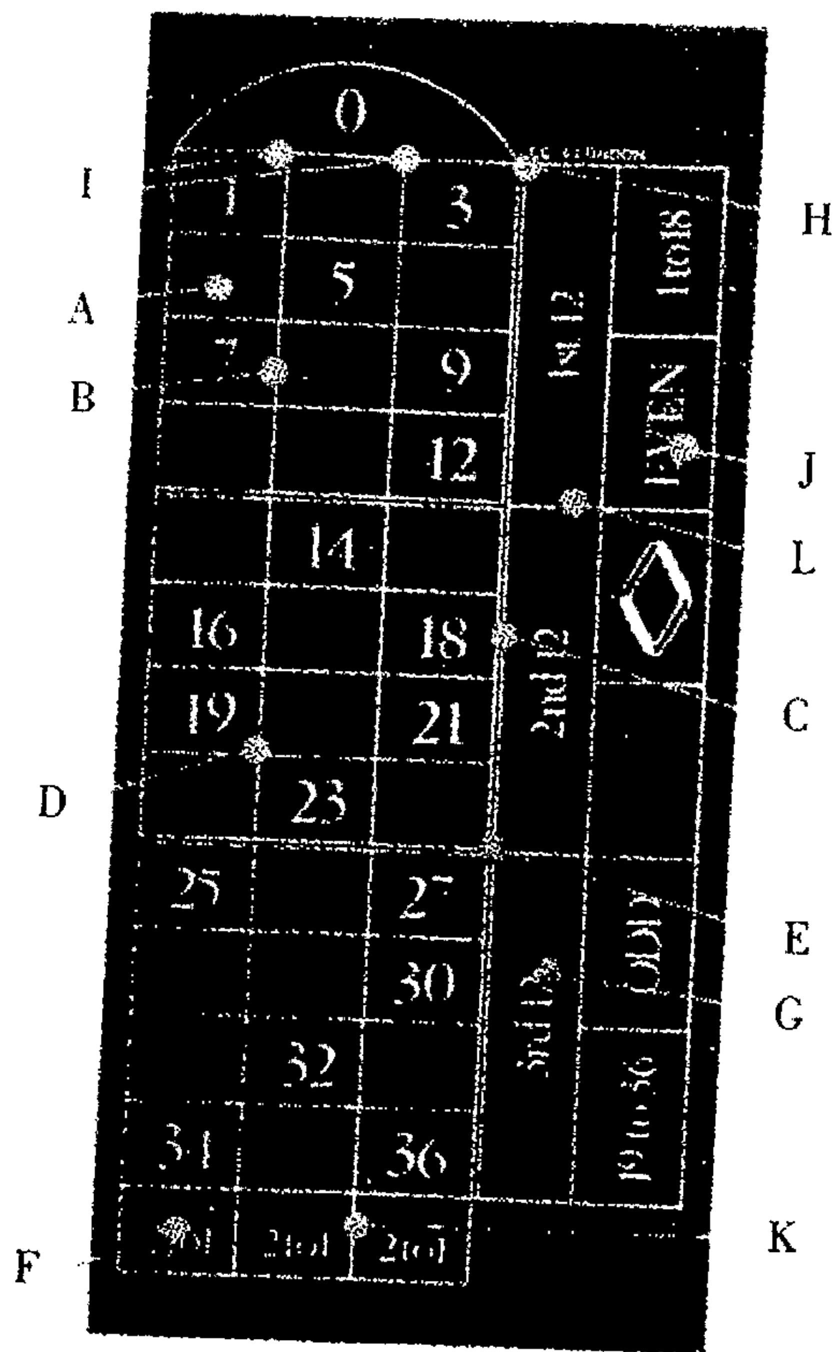
代表在 1 行数字（3 个）上下注，赔率为 1 赔 11；D 代表在 4 个数字上下注，赔率为 1 赔 8；E 代表在两行数字（6 个）上下注，赔率为 1 赔 5；F 代表在 1 列数字（12 个）上下注，赔率为 1 赔 2；G 代表在 12 个数字（1 ~ 12、13 ~ 24 或 25 ~ 36）上下注，赔率为 1 赔 2；H 代表在 4 个数字（0、1、2、3）上下注，赔率为 1 赔 8；I 代表在 3 个数字（0、1、2 或 0、2、3）上下注，赔率为 1 赔 11；J 代表在偶数数字上下注，赔率为 1 赔 1（当然也可以在奇数数字上下注，即图中标有单词“ODD”的位置，赔率也是 1 赔 1）；K 代表在两列数字（24 个）上下注，赔率为 2 赔 1；L 代表在两个区的数字（1 ~ 24 或 13 ~ 36）上下注，赔率为 2 赔 1。此外，两个菱形代表在所有红

色（或黑色）数字上下注，赔率为1赔1。我们可以假定全部37个数字出现的概率相同，因为赌场会努力完成这个目标。如果某些数字出现的概率偏大（或偏小），对赌场没什么好处，玩家确有可能乘机获利。下面我们研究一下赌场在轮盘赌中的利益。

187 在轮盘赌中，除非出现平局，玩家每投入37枚筹码，平均将赢回36枚筹码。无论你如何下注，这个平均值是不变的。所以，在每37枚筹码中，赌场将赢得1枚筹码，赌场的利润率为2.7%。

除了赌某个具体的数字，玩家还可以赌出现的数字的颜色。在这种情况下赌场的利润率减半。轮盘赌的规则规定，如果玩家赌出现的数字为黑色（或红色），而最终结果是绿色（即数字0），则为平局，赌场吃进赌注的一半，另一半退还玩家。当玩家赌数字的颜色时，赌场的利润率为1.35%。下面我们证明这个结论。

根据轮盘赌的规则，如果玩家下注2枚筹码赌红色数字出现，结果可能输掉这2枚筹码，也可能把这2枚筹码变成4枚。此外，还有1/37的概率遇到绿色数字，此时2枚筹码将变为1枚。因此，玩家的回报有3种可能为0、1或4。直接计算这种情况有点复杂，所以我们先比较另外一种情况。假设我们把轮盘赌的规则改为这样：如果最终出现绿色数字，则玩家有1/4的机会收回4枚筹码，而有3/4的机会一无所获。（1/4和3/4可以这样确定：当出现绿色数字时，玩家抛两枚硬币，如果两枚硬币都是正面，则玩家收回4枚筹码，否则玩家输掉筹码。）这种规则不会改变玩家在游戏中的收益额，但计算起来非常方便，因为此时玩家的回报只有两种可能：0或4。出现绿色数字的概率是1/37，在这1/37的情况下，玩家的回报



THE TABLE

A. STRAIGHT UP	35 to 1	G. DOZEN	2 to 1
B. SPLIT	17 to 1	H. 1ST FOUR	8 to 1
C. STREET	11 to 1	I. 3 NUMBERS	11 to 1
D. CORNER	8 to 1	J. EVEN CHANCES	1 to 1
E. SIX LINE	5 to 1	K. SPLIT COLUMN	1 to 2
F. COLUMN	2 to 1	L. SPLIT DOZEN	1 to 2

图 10.2:轮盘赌的下注区

在 1/4 的概率下为 4，在另外 3/4 的概率下为 0。综合考虑的结果是，玩家的回报在 73/148 的概率下为 4，在 75/148 的概率下为 0。1.35%是这两个值的差。我们的假设不影响游戏的结果，但可以使计算简化。

完成一个目标

保罗想去现场看欧洲杯球赛，他需要 216 英镑买机票，但他只有 108 英镑。对于保罗而言，108 和 0 没什么差别。如果他有 216 英镑，那就是天堂，否则就是地狱——即使他有 215 英镑。他是否可以到赌场碰碰运气？

赌徒间流传一个原则：如果你不得不参加一场对你不利的赌博，最好的策略是孤注一掷。对于保罗而言，我们建议他把平局当做失败处理，因为输掉一半财产和输掉全部财产没有差别。所以，在游戏中赌场的利润率始终为 2.7%。

最利落的赌法是把 108 英镑全压在红色上。他有可能立刻得到 216 英镑，概率是 18/37，即 48.6%。当然，他也可能立刻输掉所有钱，回家看电视转播。

另外，保罗也可以这样赌：把 108 英镑分成 18 份，每份 6 英镑。每次用 6 英镑赌一个数字，连赌 18 次，期待获胜 1 次。赌场的赔率是 1 赔 35，所以保罗只要赢 1 次就够了。保罗的机会有多大呢？每 1 次保罗失败的概率为 36/37，连续 18 次都失败的概率是 36/37 的18 次方，这个值接近于 0.61。因此保罗的成功率是 39%，还不如全部赌红色。

188

如果保罗决定赌一个具体的数字，他可以改进下注的方案。第一次只赌 4 英镑，因为他原来就有 108 英镑，所以只用

4英镑下注就足够了。(注意，不能只赌3英镑，否则即使获胜他的钱也不够。)实际上，只要保罗手中的赌本不少于76英镑，他就可以每次下注4英镑；当他手中的赌本不足76英镑时，他开始每次下注5英镑：当他手中的赌本不足41英镑时，他开始每次下注6英镑。利用这种下注方案，保罗可以赌22次，在这22次之中保罗只需获胜1次。22次全部失败的概率是36/37的22次方，即54.7%，所以保罗的成功率为45.3%。如果全部失败，保罗手中还剩下1英镑。保罗可以赌一种1赔5的情况试试运气（即在图10.2中类似于E的位置上下注)，如果成功，就拿刚刚到手的6英镑再赌一个数字。这样，保罗的成功率可以提高到45.7%。比每次下注6英镑好一些，不过还是不如用108英镑全买红色。

在所有可行的策略中，最有利的还是用108英镑全买红色。事实上，当遇到绿色数字时，保罗还能剩下54英镑，他可以用这54英镑继续赌。把这种因素考虑进来，保罗的成功率高达49.3%。这样他在赌场中玩的时间很短，但是他的兴趣不在于赌博，而在于欧洲杯。

我们选择108英镑作为研究的起点只是为了计算方便，不过在其他情况下我们的结论通常也是有效的：最有利的赌法就是赌博次数最少的赌法。假设保罗最初的赌本不是108英镑，而是24英镑，他的目标还是216英镑。此时他的机会小了很多，不过最优策略是类似的。他最好用所有的赌本赌一个赔率为1赔8的情况（即在图10.2中类似于D的位置上下注)。具体地说，他可以用24英镑赌4个数字，只要最终出现的数字是这4个数字中的一个他就获得成功，成功率为4/37，接近于11%。另外，他也可以把24英镑平分成两份，赌两次赔率为1赔17的情况（即在图10.2中类似于B的位置上下注)，此时

他的成功率只有10.5%。总之，最优策略是一赌定乾坤。

当保罗的赌本是24英镑时，保罗也可以这样赌：用所有筹码赌红色，如果获胜，再用所有筹码赌红色，直到赌本达到108英镑。在赌本不低于108英镑时，用108英镑赌红色，否则就把全部赌本押上。如果使用这种策略，必须期待获胜很多次，所以成功率比较低。相比之下，不如一次赌4个数字。如果保罗最初的赌本是54英镑，他可以考虑两个方案：把54英镑平分为9份，同时押9个数字；或者用所有赌本赌红色，一旦获胜就再用所有赌本赌红色。由于54已经与216的目标很接近了，所以这两种赌法没有差别，成功率都是24%。

189

赢钱的概率是多少？

假设另一个人——马歇尔——也带着108英镑来玩轮盘赌，但他的目标不是凑足216英镑，而纯粹是消遣。他将在赌场玩到6点钟，除非在此之前输光。当他离开赌场时，他赢钱的概率是多少？

答案取决于他每次的下注额和下注位置。为简单起见，我们假设他每次下注1英镑，而且只在数字区下注（即不赌数字的颜色和单双）。根据下注位置的不同，他押的数字个数m是不同的，m可以是1、2、3、4、6、12或24。（不熟悉轮盘赌规则的读者可能不容易理解，请仔细研究图10.2及文字说明。在位置A下注则m=1，在位置B下注则m=2，在位置C下注则m=3，在位置L或K下注则m=24等等。）无论他如何下注，平均每一轮他将输1/37英镑。玩370轮之后，平均输10英镑。

在他每次都押单个数字的情况下，结果的变异性最强；在他每次都押24个数字的情况下，结果的变异性最弱。变异性的强弱表示实际的结果明显偏离平均值的概率的大小。由于平均而言他将输钱，所以为了使他最终赢钱的概率较高，他应当选择变异性最强的赌法——每次都押单个数字。当然，在这种赌法下他损失较大的概率也增加了，不过我们不关心这个问题。毕竟他的目标只是使赢钱的概率尽量高，至于具体赢多少钱，他不在乎。

现在马歇尔开始玩轮盘赌，每次下注1英镑，押单个数字。刚开始时情况很可能这样：他一直输，但数额始终在35英镑以内，突然他赢了一次，赢进35英镑，总体上他赢钱了。事实上，在前35轮中他只要赢1次就可以保证赢钱。每1轮中他的胜率为1/37，即失利的概率为36/37。连续35轮都失利的概率为36/37的35次方，约等于0.383 3。所以在第35轮结束时，他有62%的概率赢钱。

但是在第38轮结束时，他赢钱的概率就不乐观了。在37轮中他投入的赌注为37英镑，为了保证赢钱他至少需要赢两次。在37轮中，他全部失利的概率是 $(36/37)^{37}$，而恰好赢1次的概率为 $(36/37)^{36}$，所以他至少赢两次的概率为 $1-(36/37)^{37}-(36/37)^{36}$，约等于27%。但是我们必须注意另一个方面：在第71轮结束时，同样赢两次就可以保证在整体上赢钱。在71轮中至少赢两次的概率为58%。 190

同样，在第73轮结束时，他赢钱的概率又变得很小。为了赢钱他至少要赢3次，这个概率只有32%；而在第107轮结束时，同样赢3次就可以保证赢钱，而此时至少赢3次的概率高很多。在第109轮结束时，他赢钱的概率又变得很小。（我非常小心地避免了对第36、72和108轮下结论。当已进行的

轮数是36的整数倍时，也可能出现收支平衡的情况，而一旦出现这种情况，则相当于游戏重新开始。）

现在问题已经很清楚了。随着游戏的进行，马歇尔赢钱的概率值呈锯齿状变化。刚开始这个概率值稳步上升，到第35轮结束时达到一个高峰，而在第37轮结束时这个概率突然变得非常低；此后这个概率值再次稳步上升，到第71轮结束时又达到一个高峰，而在第73轮结束时这个概率突然变得非常低，等等。在第36、72、108轮前后这个值的变化很剧烈。每36轮构成1个周期。此外，随着游戏的进行，每个锯齿的高度逐步递减，第35轮结束时的概率值高于第71轮结束时的概率值，第71轮结束时的概率值又高于第107轮结束时的概率值，等等。当6点钟到来时，马歇尔退出赌场，此时他赢钱的概率取决于他正处于1个周期中的什么位置。如果此时他下注的次数恰好低于36的倍数，则他赢钱的概率相当高；如果此时他下注的次数恰好高于36的倍数，则他赢钱的概率相当低。假设他每小时下注90次，两个小时内下注180次。在第179轮结束时，他只要获胜5次就可以保证赢钱，这个概率超过53%；但是在181轮结束时，要想赢钱需要至少获胜6次，这个概率只有37%。

如果马歇尔的目的就在于使赢钱的概率最大，他应当忽略具体的时间，而仅仅关注他已下注多少次。最好的情况是恰好下注35次，此时可以保证赢钱的概率最高：62%。如果他准备多玩一些时间，应当使自己下注的次数尽量接近36的倍数，但不要超过。在我们的假设之下，他下注的次数在180左右，最好的策略是在第179轮结束时退出。 191

不过这个结论中有一点纰漏。因为马歇尔只有108英镑的赌本，有可能在180轮之前他就输光了。把这个因素考虑进

来，他在第179轮结束时赢钱的概率要比我们的计算结果低0.5%。不过在第179轮结束时退出远远优于在第181轮结束时退出，这个结论仍然成立。如果马歇尔最初的赌本不是108英镑，而是181英镑，则我们的结论不用做任何修改。从另一个方面看，如果马歇尔最初的赌本是54英镑，则他在108轮之内输光的概率接近1/4，此时他的策略应做调整。

此时每次都押1个数字已不是明智之举，他应当尝试 m = 2、3、4、6或12的情况。在这些情况下，他赢钱的概率随着游戏的进行仍然呈锯齿状变化，不过变化的周期变短了。当 m = 1时，周期为36；而当 m = 2、3、4、6或12时，周期分别为18、12、9、6或3，与 m = 1时的情况相比，概率值的变化范围较小，变化也不剧烈。假设他每次都押1行数字（3个），赌场的赔率为1赔11。在第179轮结束时，他赢钱的概率为49%；而在第181轮结束时，他赢钱的概率为40%。假设他每次都押1列数字（12个），赌场的赔率为1赔2。在第179轮结束时，他赢钱的概率为40.5%；而在第181轮结束时，他赢钱的概率为38.5%。

如果马歇尔换一种赌法，每次都赌数字为红色，他赢钱的概率是否会上升？除非他准备玩几千轮，否则这种选择是错误的。无论他玩多少轮，他赢钱的概率都低于50%。在第35轮结束时，他赢钱的概率为47%；在第179轮结束时，他赢钱的概率为44%。平均而言，这种赌法赢钱的概率是最高的，但是如果采取其他赌法，他可以选择退出的时机。比如，如果他每次都押单个数字，他可以选择在第179轮结束时退出，此时赢钱的概率为53%。

必胜的幻想

如果赌场不规定下注的上限，而且允许玩家赊账，我可以提供一种必胜的赌博方案。你只需要这样做：在轮盘赌中每次都押红色数字，如果输了就把赌注加倍，直到第一次获胜为止；一旦获胜则把赌注恢复到最初的数额，重新开始。利用这个策略，无论你已经输了多少钱，只要获胜一次就可以收回所有损失，并且赢回一笔钱，这笔钱的数额与你最初的赌注相等。

不过这个策略依赖于两个不成立的假定：

○ 赌场不允许赊账。如果你利用这个策略玩轮盘赌，很可能在几轮之内输光所有赌本，此后无法继续下注；

○ 赌场有关于最大下注额的规定。即使你有充足的赌本，可是关于下注的上限的规定可能使得这个策略无法贯彻。

假设赌场规定下注的上限是 100 枚筹码，而你的目标是利用我提供的策略赢得 1 枚筹码。你成功的机会很大。你可以利用这个策略连续下注 7 次，只有在这 7 次全部失利的情况下你才可能失败，这个概率是 $(19/37)^7$，低于 1%。只要在这 7 次中有 1 次成功就可以完成目标，成功率超过 99%。

不过别高兴得太早。虽然你的成功率高于 99%，但即使成功了你的收入也不过是 1 枚筹码。但如果你失败，你的损失要大得多——你的损失是连续 7 轮的赌注，总和为 127 枚筹码。虽然失败的概率不足 1%，但平均而言你的收入还是负数。事实上，无论你用什么策略下注，结果都对赌场有利。许多赌徒都幻想发现一种必胜的策略，但实际上除非你发现了赌

场的轮盘的缺陷，否则任何策略都无法奏效。在格拉哈姆·格林尼的著作《输家通吃》中有一些精妙的例子。在这本书里，一个冒牌的数学家发明了必胜的赌博策略，当然，这个策略是错误的。

如何输钱

多数赌徒在赌场中输钱的经验极其丰富，不需要别人教他如何输钱。不过有些人进入赌场时的目标就是输掉一笔数额确定的钱。多数人参与赌博的目的是赢钱，不过结局往往是悲剧。做一个乐观的赌徒是很危险的，所以我建议你在开始赌博之前做好最坏的打算。最好的策略是事先规定好自己可以忍受的最大损失，一旦你的损失达到这个规定值，就退出赌场。这种策略的好处是你可以控制自己的损失，不必面对赌博造成的贫困。

193

下面介绍一种操作方便的策略，这种策略被称作拉勃切方法。假定你可以忍受的最大损失是 45 枚筹码，你需要制作一张赌注表。赌注表的制作和使用如下：在一张纸条的中间依次写上从 1 到 9 的九个数字，每次下注时选择两端的两个数字之和为下注额。如果这次下注失利，则把两端的两个数字勾掉；如果这次下注盈利，则不勾掉任何数字，相反，在表的末端添加你刚刚赢得的数额。每次都押红色数字。（当然你也可以押黑色数字）第一次下注时表两端的数字是 1 和 9，所以你的下注额为 10。如果第一次下注失利，则勾掉 1 和 9，第二次下注时表两端的数字为 2 和 8，所以第二次的下注额也是 10；相反，如果第一次下注盈利，则在表的末端添加数字 10（因为

你刚刚赢得 10 枚筹码），第二次下注时表两端的数字为 1 和 10，所以第二次的下注额是 11。

按照这个办法不断地在表中勾掉或添加数字，每次取两端的数字决定下注额。如果表中只剩下 1 个数字，则以这个数字为下注额。如果表中所有的数字都已被勾掉，恭喜你，你在赌场中的任务已经完成，你已经输掉了 45 枚筹码。

在你实际应用这个策略时，也许（仅仅是也许）你会发现输光这 45 枚筹码是一件很困难的事，因为赢得太频繁了。有时候你可能发现我介绍的下注方法已不再可行，因为表中两端的数字之和已超过了赌场规定的下注额的上限。这些情况只在你的运气非常好的时候发生。不过，你随时都可以退出赌场，即使你还没有完成输光 45 枚筹码的目标。少输一点钱总不是什么坏事。如果你给自己规定的最大损失不是 45 枚筹码，而是另外一个数额，你可以把从 1 到 9 的数字换成其他值，只要这些数字的总和是你规定的数额即可，勾掉或添加数字的方法不变。

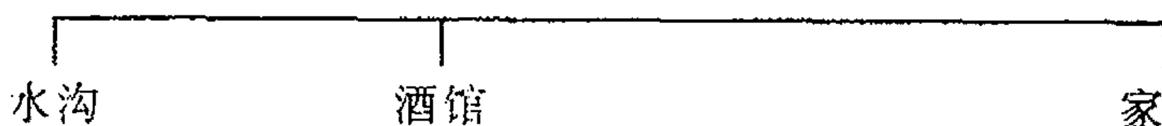
需要小心，不要把勾掉或添加数字的场合弄反了。如果你在输钱时添加数字，而在赢钱时勾掉数字，你有可能输掉一笔远远超出预算的钱。

赌徒的运气与酒鬼回家问题

赌徒的运气与酒鬼回家问题中的情况非常相似。如图所示，酒鬼从酒馆中出来，跌跌撞撞地在街上走。如果走到街右端，则安全到家；如果走到街左端，则掉进水沟。每一步他随机地向右或向左走，当然向右或向左的概率可能有微弱的差

别。每一步向右或向左与上一步的方向无关。最后他安全到家或掉进水沟是随机事件。

194



假定他每一步向右或向左的概率相等，酒馆距水沟 50 步，距家 150 步。这个问题的结论如下：

○ 掉进水沟的概率和安全到家的概率分别为 3/4 和 1/4。概率与距离成反比。

○ 平均他将在街上走 50×150=7 500 步。在一般情况下，他在街上走的步数等于两个距离的乘积。

方框中是解法的细节。请有兴趣的读者仔细研究。

酒鬼回家问题的结论可以帮助我们分析赌徒在轮盘赌中的运气。假定轮盘赌对赌徒是公平的，即赌徒每次的投入产出比为 1:1，则问题完全转化为酒鬼回家问题。酒馆的位置相当于你最初的赌本，水沟的位置相当于你输光所有赌本，家的位置相当于你进入赌场时的预期目标（假定你的计划是当你手中的筹码达到某个预定值时退出赌博）。每一轮中你胜负的概率都是 1/2。酒鬼的平均步长相当于你在每一轮中的平均赌注，酒鬼一共走了多少步相当于你在赌场中一共下注多少轮。

你的预期目标越高，则你完成目标的机会越小。随着你的预期目标的上升，你成功的概率递减，而且其极限值为 0。这说明，如果你不预先给自己规定一个退出赌场的成功条件，你将不可避免地在赌场中输光，即使轮盘赌对于你和赌场同样有利（事实上在轮盘赌中赌场总是略微有利一些）。

如果你把自己在每一轮中的平均赌注增加一倍，结果如

何？这种情况相当于在酒鬼回家问题中酒鬼把自己的平均步长增加一倍。问题等同于一切情况不变但街长缩短了一半。你成功或失败的概率不受影响，当时你在赌场中花费的时间将缩短很多。你参赌的平均轮数降低为原来的 1/4。

195

酒鬼回家问题

假定酒鬼迈出的每一步的步长相等，向左走的概率与向右走的概率相等。街长为 L 步，水沟的坐标为 0，家的坐标为 L，酒馆的坐标为 n，于是酒馆距水沟和家的距离分别为 n 步和 L－n 步。

当酒鬼距水沟 k 步（即当前的坐标为 k）时，用 p（k）表示在这种情况下他安全到家的概率。当 k＝0 时，酒鬼已经掉进水沟，他已不可能安全到家，所以 p（0）＝0；当 k＝L 时，酒鬼已经安全到家，所以 p（L）＝0。当酒鬼距水沟 k 步时，他有 1/2 的概率向左走（距水沟的距离变成 k－1），另外有 1/2 的概率向右走（距水沟的距离变成 k＋1），所以

$$p(k) = 1/2 \times p(k+1) + 1/2 \times p(k-1)$$

当 k 取任意值的情况下，这个公式成立。在数学上这个公式称为“差分方程”，数学家已经发明了处理这类问题的通用方法。最后的解是 p（k）＝k/L。结论：

如果最初酒馆与家的距离为 n 步，则安全到家的概率为 n/L。

利用类似的方法可以求出酒鬼平均走了多少步。以 k 表示酒鬼的当前坐标，以 T（k）表示在此情况下酒鬼平均走了多少步。当 k＝0 或 L 时，酒鬼不会再走，所以 T（0）＝T（L）＝0。依据同样的推理，可以得到

$$T(k) = 1 + 1/2 \times T(k+1) + 1/2 \times T(k-1)$$

这也是一个差分方程，它的解是 T（k）＝k×（L－k）。结论：

如果最初酒馆与家的距离为 n 步，则酒鬼走的步数为 n（L－n）。

赌场不提供对双方公平的游戏。不过你可以和一个朋友设计一个对你们两人公平的游戏。假定你们两个用抛硬币的方法赌钱，你的赌本为 1 英镑，而他的赌本为 1 000 英镑，每 1 轮的赌注额为 1 英镑，游戏结束的条件为某一方输光。你在游戏中的目标是把自己的资金从 1 英镑变成 1 001 英镑，你的成功率为 1/1 001。游戏平均将进行的轮数为双方赌本的乘积，所以平均你们将抛 1 000 次硬币。不过另一方面，游戏仅进行 1 轮就分出胜负的概率高达 1/2。

这说明，游戏实际进行的轮数可能与平均值有很大差距。如果你利用平均值估计游戏将进行多少轮，出错的可能性很大。

以上分析可以帮助你估计自己在赌博中的机会。在一场公平游戏中，如果你的目标是把自己的资金提高到原来的 10 倍，则成功率只有 1/10。实际情况是在赌场中游戏总是对你不利，所以你的成功率比 1/10 还要低。 196

下面我们研究轮盘赌中实际发生的情况。假定一个赌徒每次都押红色数字，则他盈利的概率与失利的概率不完全相等。这相当于在酒鬼回家问题中酒鬼向左走和向右走的概率不相等的情况（向左走的概率略大）。问题的解法没有根本性的变化。请研究下面方框中的文字。

轮盘赌中的实际情况

假定每次赌徒都押红色数字。根据我们以前得到的结论，他盈利的概率为73/148，失利的概率为75/148。在某些国家，赌场规定出现绿色数字（即0）是赌场完全吃进赌注，此时以上两个概率值应调整为18/37和19/37。

用p表示赌徒盈利的概率，用q表示赌徒失利的概率，q=1-p。用x表示q/p，用y表示q-p。由于赌博总是对赌场有利，所以x大于1，y大于0。利用我们在酒鬼回家问题中使用的方法可以得到两个类似的差分方程，解这两个方程可以得到结论：

赌徒成功的概率为$\frac{x^n-1}{x^L-1}$

赌徒的平均下注次数为$\frac{n}{y}-\frac{L}{y}\times\left(\frac{x^n-1}{x^L-1}\right)$

在英国赌场规则中，p=73/148，q=75/148，于是x=75/73，y=1/74。如果赌徒目标为218英镑，每次下注1英镑，每197次都押红色数字，根据其赌本的不同，结果如表10.1所示。

表10.1

赌徒的赌本	54	90	108	144	162	180	198
成功的概率（%）	1	3	5	14	23	38	61
平均下注次数	3 800	6 200	7 200	8 400	8 300	7 300	4 850

如果赌徒最初的赌本为108英镑，每次谨小慎微地下注1

英镑，则他最后成功的概率仅有 5%。我们以前讨论过，如果赌徒豪赌一次，把 108 英镑同时押在红色数字上，则成功率高达 49%。胆小的赌徒通常是不幸的赌徒。赌场就是利用这种情况盈利的：如果赌徒每次下注 1 英镑，赌场可以利用漫长的 7 200 轮耐心地吃掉全部的 108 英镑的赌本。你的平均下注额越低，你下注的次数就越多；而你下注的次数越多，赌场在游戏中的优势越能充分地发挥作用。

赌博中的物理学

在轮盘和小球开始旋转以后，小球在哪个数字上停下来是由物理学规律决定的。有些赌徒一直梦想发明一种在赌场中必胜的方法——利用计算机精确分析轮盘和小球的运动状态，预测将出现的数字。大致的设想是把轮盘上的数字分为若干区域，让计算机判断小球最终落到哪个区域的概率较高。比如说，计算机预测小球将落在数字 10 附近，赌徒则在 10 附近的几个数字（24，5，10，23 和 8）上下注。

赌场禁止顾客把计算机带入赌场，另外在小球的速度慢下来以前就宣布"停止下注"。在每一轮中，赌场的额外优势相当小，如果顾客可以预测小球落在某些区域的概率略高（或略低），赌场很容易亏本。比如，如果顾客发现小球落在｛0，32，15，19｝的区域内的概率比平均值低一半，他就可以把赌注押在另外 33 个数字上，他可以保证 3.1%的利润率。

不要低估物理学计算的难度。托马斯·柏斯写过一本名为《赌场中的力学》的书，书中介绍了类似的尝试。你需要计算轮盘和小球的速度，并考虑摩擦力的影响。初始状态的一点非 198

常微弱的差别将使最后的结果迥然不同，不过顾客不需要得到精确的结论。如果顾客拥有充足的资金和耐心，他只要某些区域对应的概率偏高或偏低就足以制胜。

如果你真想这样干，千万要小心——你可能被赌场列上不受欢迎的黑名单。在《赌场中的力学》一书中，一群大学生把计算机藏在鞋里带进赌场，偷偷地用脚趾按键盘。他们的计划破产了，因为赌场识破了他们的计谋。不过他们在其中获得了乐趣。

从赌场的角度看

对于赌场而言，最理想的情况是每个数字上都有人下注，而且每个数字上的赌注大致相等。这样，无论最后出现的是什么数字，赌场都可以保证每轮有 2.7%的收入。最可怕的情况在于，一个顾客在几个数字上下重注，而且获胜。一旦出现这种情况，赌场可能一下子输掉一个月的利润。为了防备这种情况发生，赌场需要储备一大笔钱作为预备金，预备金的数额取决于赌场允许的最大赌注额。预备金是赌场的安全保障。

规定每轮的最大下注额实际上为赌场设定了每轮中的最大损失的上限。赌徒们可能勾结起来，每个人都在同一个数字上下最大赌注，这种情况是赌场禁止的。正如保险公司要求客户不得隐瞒必要的信息，赌场要求顾客遵守规则，不许串通，不许玩赖。

如果赌场的轮盘不够规范，赌场有可能因此蒙受最大损失。轮盘是一种高精度仪器，需要定期地监测、清理和维护，并且经常调换各个赌桌上的轮盘，以保证每一轮的结果都是不

可预测的。如果你确知某个赌场在这方面疏于防范，你就有机可乘。当然你需要仔细观察或者买情报。

每张赌台上的庄家都是精于计算的好手。每一轮结果出现以后，庄家立刻就能算出应当赔给哪些玩家多少钱。这种计算对于他们太熟悉了，多数情况下他们根本不假思索，直接就把筹码推到玩家面前，数额当然是正确的。

虽然赌数字的颜色对玩家更有利，但实际上这种下注方法并不流行。假设你有 18 枚筹码，用这些筹码押红色数字比押 18 个具体的数字更有利，不过多数赌徒宁愿在具体数字上下注。这种现象可以从赌徒的心理上解释：押数字颜色的赔率是 1:1，很难让人振奋；而押具体数字的赔率高很多（最高达到 1:35），这种情况才叫刺激。

如果赌场可以自由地做广告，他们可以通过一个对比展示轮盘赌的优势。如果一个人只有 1 英镑，他如何能成为百万富翁呢？在轮盘赌上连续赌 4 轮，每次都押 1 个数字，押上当前的所有赌注。如果 4 轮都获胜，则他已经拥有 170 万英镑的财产。当然，这种事发生的概率不大，只有 190 万分之一。你觉得这个概率太小吗？可是比较一下大英帝国彩票中的中奖概率，在大英帝国彩票中通过 1 英镑达到同等成就的概率只有这个概率值的七分之一。

陀思妥耶夫斯基的赌博生涯

陀思妥耶夫斯基（俄国著名作家）曾经狂热地迷恋轮盘赌，尤其是在 1863 ~ 1871 年他在德国疗养的时候。1866 年他完成了《赌徒》一书，这本书带有很强的自传色彩。在那段时

间，他经常在书信里宣称他已经掌握了在轮盘赌中制胜的秘诀。他坚信："只要严格地按照此秘诀行事，失败是不可能的。"类似的话他重复过许多次。1867年5月，他在汉堡再次表达充沛的信心："在赌博中极端的冷静和谨慎可以保证必胜，你想赢多少钱就可以赢多少钱。但是不能急于求成，你可能需要赌很长时间，持续很多天，在运气不好时也要坚持。""冷静而周密的计算可以保证必胜。""不进行计算而仅仅依靠运气的赌徒是傻瓜。""如果你每天都赌，但每次只赌一点点，你没有获胜的机会。"

为了证明自己的信念，他进行了坚韧不拔的努力，结果每每陷入赤贫的窘境。但是他拒绝相信他的诀窍是错误的。他宣称，自己在实战中太兴奋了，偏离了必胜的策略。他在这段时间写的信经常是借钱的。他最后一次玩轮盘赌可能是在1871年4月，此时他的信念已经动摇了。也许他意识到了凭借赌博已经不可能偿还赌债。另一个事件可能也是促使他退出赌博的原因：德国政府下令关闭了赌场。究竟是什么原因使他浪子回头，我们不得而知。

《赌徒》一书多次提到轮盘赌，但是在第十四章有一个小错误。根据陀思妥耶夫斯基的描述，亚历克斯（书中的主人公，也是作者的化身）在13~24这12个数字上下注，并且获胜。随后亚历克斯有一段评论："这种下注方法的赔率是1赔3，而获胜的机会是2:1。"据我所知，这个世界上还没有哪家赌场符合亚历克斯的描述。随后书中又说，亚历克斯的下注额是80，获胜之后取回200。根据这种说法，赌场的赔率应当是2赔3，而非1赔3。这是怎么回事呢？也许陀思妥耶夫斯基输钱输糊涂了。

赌场扑克

赌场扑克与在美国西部片中经常出现的扑克游戏不同。在美国西部片中，胆大心细的主人公往往用一手烂牌把对手吓倒，对手明明拿着一手好牌却缴了枪。但是在赌场扑克中，这种虚张声势的技巧派不上用处。赌场中的资金有6%流通在这种游戏中。在这种游戏中，玩家的主要技巧在于判断是否应当增加赌注，此外则完全凭借运气。一副扑克有52张牌，在其中任取5张构成一手牌，一共有$^{52}C_5 = 2\ 598\ 960$种组合。每一手牌可以归入某个级别。下面是各个级别对应的赔率，括号中的数字表示这种级别包含多少种组合：

201

同花顺（40）：5张连续的牌，同样花色。赔率50:1。

四条（624）：有4张大小相同的牌。赔率20:1。

三带二（3 744）：5张牌中有3张大小相同，另外两张大小相同。赔率8:1。

同花（5 108）：5张牌花色相同，但不连续。赔率6:1。

顺子（10 200）：5张连续的牌，但花色不完全相同。赔率4:1。

三条（54 912）：5张牌中有3张大小相同，但另外两张大小不同。赔率3:1。

两对（123 552）：5张牌中有两对大小相同的牌。赔率2:1。

对子（1 098 240）：5张牌中仅有1对大小相同的牌。赔率1:1。

AK（167 280）：5张牌中有1张A和1张K。没有对子，也不是同花或顺子。赔率1:1。

烂牌（1 135 260）：不符合以上9种条件的牌。没有赔率。

以上顺序也是判断一手牌的大小的标准，从上到下级别递减。92%的牌属于最后3个级别。如果有两手牌属于同一级别，可以通过具体的牌的大小判断两手牌的大小，比如说{8，8，8，J，J}和{6，6，6，K，K}都属于"三带二"，但前者较大。在非常偶然的情况下，会出现两手完全相同的牌，比如说两个人的牌都是{A，A，7，6，2}，这种情况称为"平局"。

庄家代表赌场与玩家对抗，每个玩家各自独立地下注。所以在你参与赌场扑克时，你可以忽略其他玩家的存在，假设只有你一个人与庄家对抗。在你看牌之前，你必须确定第一轮的下注额。在你下第一轮赌注时，只有庄家的最后一张牌是翻开的，其他牌都扣在桌面上。此后你可以看自己的牌，然后决定是否增加赌注。如果你决定不增加赌注，则你已经投降，庄家吃进你的赌注，双方都不必亮牌。

如果你决定增加赌注，你的第二轮下注额必须是第一轮下注额的二倍。此时下注过程结束，你一共下了两笔赌注，这两笔赌注的赔率分开计算。然后庄家开牌。如果庄家的牌是烂牌，则不论你手里的牌是什么你都获胜。不过在这种情况下庄家只赔你的第一轮赌注，赔率为1赔1，第二轮赌注不予考虑。这种情况发生得极为频繁，因为在所有的5张牌组合中有

202

接近44%的烂牌。也许此时你手中拿着四条，心中怀着必胜的信心和大胜一场的喜悦，可是庄家手中竟然是一副烂牌，你只赢到一笔与第一轮赌注相等的小额收入。

如果你决定增加赌注，而庄家手中不是烂牌，这时你的两轮赌注都起作用。如果庄家的牌比你大，庄家吃进你的所有赌注；如果出现平局，你取回所有赌注，不吃不赔；如果你的牌比庄家的牌大，庄家赔钱，但两笔赌注的赔率分开计算：第一轮赌注的赔率为1赔1，第二轮赌注的赔率由你的牌的级别决定。

下面是四个例子，请考虑在各种情况下的结果（假设你的第一轮赌注为1枚筹码）：

A. 你的牌是两对，你决定增加赌注。庄家的牌是｛A，Q，10，9，5｝，不是同花；

B. 你的牌是同花，你决定增加赌注。庄家的牌是对子；

C. 你的牌是两对，你决定不增加赌注。庄家的牌与A中相同；

D. 你的牌是三条，你决定增加赌注。庄家的牌是顺子。

答案如下：

A. 庄家的牌是烂牌。你赢得1枚筹码。你的第二轮赌注没起作用；

B. 庄家的牌不是烂牌，你获胜。你赢得13枚筹码。第一轮赌注赢得1枚筹码，第二轮赌注赢得12枚筹码（第二轮赌注为2枚筹码，同花的赔率为6:1）；

C. 你输掉你第一轮筹码。你不知道你可以取胜；

D. 庄家获胜。你输掉3枚筹码。

无论庄家的牌多么好，在每一局中你顶多输掉3枚筹码。相反，你最多可能赢得101枚筹码，如果你有幸拿到一副同花

顺。不过你大可不必为此兴奋，你拿到同花顺的概率非常低。如果你每分钟玩1局，每天玩1小时，每年玩250天，平均而言在4年之中你会拿到一副同花顺。而且，即使你拿到了同花顺，如果庄家的牌是烂牌，你的好牌也不起作用，庄家只赔给你1枚筹码。庄家手中不是烂牌的概率只有56%。如果你坚持玩8年，你会遇到1次赢回101枚筹码的好运气。

203

这个游戏当然是对庄家有利的。在游戏中，你需要考虑四个问题：

○ 在什么情况下你应当增加赌注？

○ 在决定是否增加赌注之前，庄家的那张已经亮开的牌是否可以帮助我做决定？

○ 在这个游戏中庄家的优势有多大？

○ 我在这个游戏中的投资回报率是多少？

如果你手中是什么牌与庄家手中是什么牌互不影响，这些问题会简单得多。假设你和庄家分别从两副牌中取牌，则你们两个人手中的牌就是相互独立的。但实际情况是，在你从一副牌中取出5张以后，这副牌只剩下47张。庄家从47张牌中取出5张，可能的组合有${}^{47}C_5 = 1\ 533\ 939$种，其中任何1种出现在庄家手中的概率相等。这些组合的总数是固定的，但是这些组合的内容受你手中的牌的影响。复杂之处在于，虽然多数情况下你手中的牌对庄家手中的牌的影响相当小，但某些情况下却是不容忽视的。

举一个例子。虽然四张A带一张K比四张Q带一张J大，但是我认为后者比前者更有利。这两手牌获胜的概率都非常大，但后者赢得41枚筹码的机会更大。为了赢得41枚筹码，庄家的牌必须不是烂牌。当你持有四张A带一张K时，庄家已经不可能持有AK，他手中的牌不是烂牌的概率低于40%。

相反，如果你手中的牌仍然是四条，但不包括 A 和 K，则庄家手中的牌不是烂牌的概率在 47% 左右。基于相同的原因，同样是同花顺，AKQJ10 不如 QJ1098 更有利。为了取得高额奖金，必须持有同花顺或四条，这种好运到来的概率只有 1/4 000。但是当好运到来的时候，如果庄家手中却是烂牌，我们将多么伤心！

为了决定是否增加赌注，我们需要对比两种决定带来的后果，选择平均收益最大的一种。（这与在西洋双陆棋决定是否接受加倍类似，我们在第五章讨论过。）如果不增加赌注，我们将输掉 1 枚筹码，所以，如果增加赌注带来的平均收益高于 −1，我们就应当增加赌注。为了判断平均收益，我们要进行耐心的计算。我们需要判断庄家手中的牌属于以下几种情况的概率：

（1）烂牌；

（2）不是烂牌，但比我们手中的牌小；

204

（3）不是烂牌，并且与我们手中的牌一样大；

（4）不是烂牌，并且比我们手中的牌大。

计算情况（3）通常很简单。四种情况对应的全部组合的总数为 ${}^{47}C_5$，如果考虑到我们可以看到庄家的 1 张牌，这个总数为 ${}^{46}C_4$。所以我们的主要任务是计算前两种情况。计算过程相当麻烦。杂色的 J10873 和杂色的 J10987 出现在庄家手中的概率相等，但是后者远比前者大。我们按照从草花 2 到黑桃 A 的顺序为 52 张牌编号，编号是 1 ~ 52。每种组合可以按赌局中的大小排序。

比较两副同属于 AK 的牌：

第一副牌：黑桃 A，红心 K，方块 J，草花 8，黑桃 2；

第二副牌：红心 A，黑桃 K，红心 J，黑桃 8，红心 2。

如果我们持有第一副牌而庄家持有第二副牌，则双方一样大，结果是平局。但是，第一副牌比第二副牌好！如果把第二副牌换成第一副牌，则情况（1）、（2）、（3）出现的概率上升，而情况（4）出现的概率下降。证明这个结论很麻烦，不过思路很简单：当我们持有第一副牌时，庄家持有同花的概率较小。无论我们持有第一副牌或第二副牌，庄家持有对子、两对、三条、四条或三带二的概率是一样的，但是当我们持有第二副牌时，庄家持有同花的概率较大。模仿乔治·奥维尔的名言表达即：第一副牌与第二副牌是平等的，但第一副牌比第二副牌更平等。[①] 方框中的文字是详细的分析。

这个例子的目的是展示计算的复杂程度。这两副牌在价值上的差别仅相当于一枚筹码的千分之一，但是这个差别只有经过仔细地运算才能发现。在实战中，当我们面对抉择时通常需要类似的计算。

在这个游戏中，虚张声势不起作用。也许你这样认为：如果当我们持某一副牌时应当增加赌注，则当我们持一副更好的牌时也应当增加赌注；如果当我们持某一副牌时不应当增加赌注，则当我们持一副更糟的牌时也不应当增加赌注。也就是说，我们可以设计一副“标准牌”，当我们的牌比标准牌好时就增加赌注，否则不增加赌注。如果不考虑我们在作决定以前能看到庄家的一张牌，这种想法是正确的。但实际情况是，在你决定是否增加赌注之前，已经有一张庄家的牌翻开了。这使得问题变得极为复杂。有时候我们手中的牌比标准牌好，但是那张翻开的牌促使我们不增加赌注；有时候情况相反，我们手

205

① 在这里作者想引用一句西方名言：每个人都是平等的，但有些人比其他人更平等。——译者注。

中的牌比标准牌糟，但是那张翻开的牌促使我们增加赌注。这说明，标准牌的确定依赖于那张翻开的牌。

首先我们不考虑那张翻开的牌。在这种情况下如何确定标准牌？一开始我们可以从一些具体的牌例出发，根据我们手中的牌判断庄家的牌比我们好或糟的概率。在某些牌中，我们应当决定增加赌注，在另外一些牌中，我们应当决定不增加赌注。标准牌应当处于这两类牌之间。不断缩小讨论的范围，最后就可以发现标准牌。现在进入具体的过程。计算过程需要谨慎，但是并不需要特别高的精确度。

假设我们已经把所有的可能组合按强弱顺序列在一张表中。如果我们此刻手中的牌处于表的上半部分，则局面对我们有利；否则局面对我们不利。标准牌应当居于表的中点附近。由于有44%的组合比 AK 更弱，我们首先从 AK 类牌中最弱的一种开始讨论，即 AK432。

不考虑我们手中的牌对庄家手中的牌的影响。我们手中的牌是 AK432，庄家手中的牌有 44%的概率是烂牌，在另外的 56%的概率下，庄家的牌比我们强。假定我们决定增加赌注。如果庄家手中的牌是烂牌，则我们赢得 1 枚筹码；否则我们输掉 3 枚筹码。此时我们的平均收益为

$$(1)\times(0.44)+(-3)\times(0.56)=-1.24$$

这个值比 -1 小 24%。所以当我们持有 AK432 时，不应当增加赌注。标准牌应当比 AK432 强。

下面讨论 AK 类牌中最强的一种，即 AKQJ9。假定我们持有 AKQJ9，并决定增加赌注。如果庄家是烂牌，则我们赢得 1 枚筹码；如果庄家是 AK，则我们赢得 3 枚筹码；如果庄家是对子或更强（概率接近 50%），则我们输掉 3 枚筹码。计算方法不变，不过我们增加一点计算精度。此时我们的平均收益为

$$(1)\times(0.437)+(3)\times(0.064)+(-3)\times(0.499)=-0.868$$

这个值比 -1 大 13%，所以此时增加赌注是正确的。

一副牌不属于 AK 类的概率为 93.6%。而当我们手中的牌不属于 AK 类时，我们的策略已经得到了。如果我们的牌比 AK 弱，则不应当增加赌注；如果我们的牌比 AK 强，则应当增加赌注。如果我们的牌比 AK432 弱但我们增加赌注，我们的损失将增加 24%；如果我们的牌比 AKQJ9 强但我们不增加赌注，我们的损失将增加 13%。注意到 24% 接近于 13% 的二倍，所以我们猜想标准牌居于 AK 类牌的前三分之一附近，大致是 AKJxy。当双方手中的牌都是 AK 时，胜负取决于对小牌的比较，所以有时候小牌是很重要的。我们尽量把小牌计算精确。

假定我们手中的牌不超过对子。庄家手中的牌一共有 N 种可能，分为四种情况：

○ 烂牌。对应 x 种可能；

○ AK 或对子，但比我们的牌弱。对应 y 种可能；

○ 与我们手中的牌一样。对应 z 种可能；

○ 比我们手中的牌强。对应 N - x - y - z 种可能。

在这四种情况中，我们的收益分别为 1，3，0，-3。因此，我们的平均收益为

$$(1)\times\frac{x}{N}+(3)\times\frac{y}{N}+(-3)\times\frac{N-x-y-z}{N}=\frac{4x+6y+3z}{N}-3$$

为了使这个值大于 -1，应满足以下条件：

$$4x+6y+3z>2N$$

所有满足这个条件的牌都应当增加赌注。我们刚才讨论过两副牌，都由 AKJ82 构成。第一副牌是“黑桃 A，红心 K，方块 J，草花 8，黑桃 2”；第二副牌是“红心 A，黑桃 K，红心

J，黑桃 8，红心 2"。非常有趣的是，第一副牌满足条件，而第二副牌不满足条件。如果我们仔细推敲所有由 AKJ82 构成的组合，会发现一个简洁的结论：一副由 AKJ82 构成的牌如果包含全部四种花色，则赢得增加赌注；如果只包含二或三种红色，则不应增加赌注。

进一步说，如果一副牌比 AKJ82 强，则应当增加赌注；如果一副牌比 AKJ82 弱，则不应增加赌注。

下面考虑庄家那张翻开的牌对我们的决定的影响。这张牌是什么有 47 种可能（因为我们已经取出了 5 张牌）。有时翻开的牌会鼓励我们增加赌注，有时则相反。缜密地分析这个问题相当困难，幸好彼德·格里芬和约翰·格温已经解决了这个问题。他们的方法是穷尽所有的可能性（一共有 210^{13}种），分别讨论 $4x+6y+3z>2N$ 这个条件是否被满足。

其中 $N={}^{46}C_4=163\ 185$

粗略地说，以下三种情况（按重要顺序排列）鼓励我们增加赌注：

○ 翻开的牌与我们手中的某一张牌一样大；

○ 翻开的牌比较小；

○ 翻开的牌与我们手中的某一张牌属于同一花色。

对于第一种情况，假设翻开的牌是一张 7，而我们手中也有一张 7，则我们手中的牌使得庄家持有一对 7 的概率下降，另外这张翻开的 7 也使得庄家的牌组成其他对子的概率下降，因此庄家的牌比我们强的概率下降。第二种情况很好理解。对于第三种情况，如果翻开的牌与我们手中的某一张牌属于同一花色，则对手持有同花的概率下降。

我们不可能深入讨论问题的细节，只提供以下五个指导原则。前三个原则是确切无疑的，后两个原则是近似结果。

（1）当你持有 AKQJx 或更强的牌时，增加赌注；

（2）当你持有烂牌时，不增加赌注；

（3）当你手中的牌是 AKJxy 时，如果翻开的牌与你手中的某张牌一样大（有 15 种可能），则增加赌注；否则不增加赌注；

（4）当你持有 AKQxy 时，你可以比持有 AKJxy 时更主动。如果翻开的牌与你手中的某张牌一样大，则增加赌注。此外，如果翻开的牌比较小，你也可以增加赌注。例如，如果你手中的牌是 AKQ94，则仅在翻开的牌是 10 或 J 时不增加赌注，在其他情况下总是增加赌注。如果你手中的牌是 AKQ54，则仅在翻开的牌在 6～J 之间时不增加赌注，在其他情况下总是增加赌注；

（5）当你手中的牌是 AK10xy 或更弱时，如果翻开的牌是 A、K 或你手中没有的牌，则不应增加赌注。此外还有一些情况使得你不应增加赌注。这说明，在全部 47 种可能中，你应当增加赌注的情况不超过 9 种。例如，如果你持有 AK932，则只有翻开的牌是 2 或 3 时你应当增加赌注，对应的情况只有 6 种。

彼德·格里芬和约翰·格温的计算精度很高，你可以充分信任他们的结论。不过你可能已经发现了，那张翻开的牌所提供的信息没有多大用处。

如果你严格遵循这五个原则，则有 48% 的牌你不会增加赌注，另外 52% 的牌你会增加赌注。在每局中你的平均赌注额为

$$(0.478) + (3) \times (0.522) = 2.044$$

彼德·格里芬和约翰·格温利用计算机模拟游戏过程，结论是如果你严格遵循指导原则，平均每局你将损失 0.05 枚筹码。

这说明，赌场的利润率为（0.05）/（2.044），约等于2.5%，和轮盘赌很接近。

通常，你手中的牌越强，则你可以期望的收益越好。当然也存在例外，比如我们前面讨论过的例子，4张Q的预期收益比4张A的预期收益高。不过这种情况我们可以不考虑。我们关心另外一些问题：当我们持什么样的牌时可以保证获胜的概率高于1/2？是否存在一副牌，使得当我们的牌比这副牌强时我们获胜的概率高于1/2，而当我们的牌比这副牌弱时我们获胜的概率低于1/2？答案是否定的，这样的牌并不存在。

当我们持有66987时，我们获胜的概率高于1/2。有趣的是，661052比66987大，但是661052获胜的概率却低于1/2。这种情况很容易理解。无论你手中的牌是66987或661052，庄家的牌都可能是对子，在两种情况下，这些对子的总数相同。然而，当你的牌是66987时，庄家的牌是一对2、3、4或5的概率较高，而当你的牌是661052时，庄家的牌是一对2或5的概率下降。这个差别使得当你持有661052时庄家获胜的概率较高。

发明这个游戏的人一定有惊人的想象力。初学者刚参与这个游戏时不必担心对手虚张声势。无论游戏如何进行，总是对庄家有利。

比九点

这是一种简单而流行的游戏，顾客的赌注有1/6押在这种游戏中。顾客和庄家每人取两张牌，牌放在赌桌上，背面向上。胜负由比较双方的点数决定。一方的点数即他的两张牌的

点数之和。A的点数是1，2~9的点数分别是2~9，10、J、Q、K的点数是0。如果一方的点数超过10，则只取个位数字作为他的点数。例如，当你持有9-6、Q-5或A-4时，你的点数都是5。点数较高的一方获胜，如果双方的点数相等，则为平局。

玩家可以选择是否要第三张牌，但是有严格的规定。如果你的点数不超过4，你必须要第三张牌；如果你的点数不低于6，你必须不要第三张牌；只有当你的点数恰好为5时，你有权自由决定是否要第三张牌。这种规定的目的是保护其他顾客的利益，因为其他顾客可以在双方中的任何一方上下注。第三张牌在桌面上亮开。

当一个玩家的头两张牌的点数为5时，他可以自由地决定是否要第三张牌，但是最佳策略是随机地决定是否要第三张牌，在9/11的概率下决定要第三张牌，而在2/11的概率下决定不要第三张牌。利用这种策略，可以保证庄家的平均利润率只有1.28%。这个值明显低于轮盘赌和赌场扑克中的庄家利润率。

临场时如何确定当前的局面属于9/11或2/11呢？你可以随身带一副备用的扑克牌，去掉其中的K和Q，只留下44张牌。[①] 当你面临选择时，随机地抽一张牌，如果抽出的牌是A或4，则决定不要第三张牌，否则要第三张牌。选择哪两张牌作为对于“决定不要第三张牌”的提示，完全取决于你与自己的约定，你应当小心不要让庄家洞悉你的心思，另外你最好不时地改变你的约定。最后需要注意的是，你应当在每局开始之前都这样抽一张牌，如果这一局中你的头两张牌的点数果然是

① 西方人的扑克中不包括王牌。——译者注。

5，则你已经有了决定，如果不是5这样做也没什么坏处。如果你只是在看到自己的点数是5时才抽一张牌，庄家就可以猜到你最初的两张牌的点数是5，你的行为就泄密了。

下面我们继续描述这种游戏。首先由玩家决定是否要第三张牌。此后庄家也要决定是否要第三张牌。赌场事先已经为庄家确定了最佳策略，庄家只需要根据具体情况作出判断。庄家的点数有八种可能：从0到7。（赌场规定，如果点数是8或9，则不取第三张牌。）从庄家的角度看，玩家的行为有11种可能：玩家决定不要第三张牌；玩家决定要第三张牌并得到一张点数为0的牌；玩家决定要第三张牌并得到一张点数为1的牌；……玩家决定要第三张牌并得到一张点数为9的牌。因此，庄家所面对的局面有$8 \times 11 = 88$种可能。赌场为庄家预备的策略相当于一张表，这张表规定了庄家面对88种局面中的任何一种时应如何抉择。所有的可能策略有多少种呢？每一种局面下有两种可能策略："要"或"不要"，所以总共有2^{88}种可能策略。这个数相当巨大，是一个长达27位的十进制数。

赌场需要在2^{88}种可能策略中选出最佳策略。这个任务粗看起来很困难，不过我们可以利用第六章中介绍过的"优先策略"的概念。某一类策略明显比另外一类策略强。比如说，当庄家的当前点数为0时，毫无疑问他应当要第三张牌。这样一下子就淘汰了所有可能策略中的一半。反复利用这个思路，绝大多数局面是可以判定的。最后只剩下四种不容易判定的局面：

（1）庄家的点数为6，玩家不要第三张牌；

（2）庄家的点数为5，玩家的第三张牌是4；

（3）庄家的点数为4，玩家的第三张牌是1；

（4）庄家的点数为3，玩家的第三张牌是9。

211 现在庄家的可能策略只剩下 $2^4 = 16$ 种。利用优先策略的概念比较这 16 种策略，又可以淘汰其中的 11 种，庄家的可能策略只剩下 5 种。

问题已经简化到这种情况：庄家和玩家对抗，庄家有 5 种策略（即面对以上 4 种局面时决定要或不要第三张牌），而玩家有两种策略（即当玩家最初的点数为 5 时玩家决定要或不要第三张牌）。回忆一下我们在第六章中得到的结论，庄家的最佳策略是由两种策略构成的混合策略，这两种策略是“不要、要、不要、要”和“要、要、不要、要”。请注意，庄家在面对局面（2）、（3）、（4）时的策略已经确定了，一定是“要、不要、要”。最后一个问题是确定局面（1）之下的策略。

总之，在全部 88 种可能局面下，有 87 种局面的最佳策略已经确定，惟一需要讨论的是“庄家的点数为 6，而玩家不要第三张牌”时的情况。在这种局面下，庄家和玩家各有两种策略，一共可以构成 4 种组合。赌场在发牌时通常把 6 副扑克牌混在一起，如果假设每张牌被抽出的概率相等，则可以确定庄家在各种组合下的收益情况。利用这些结果可以形成支付矩阵：

	庄家在局面（1）下“要”	庄家在局面（1）下“不要”
玩家在点数为 5 时“要”	-4121n	-3705n
玩家在点数为 5 时“不要”	-2692n	-4564n

表中 n = 16/4 826 809。这些数字是通过繁琐的过程求得的平均值。

约翰·肯曼尼和拉里·斯奈尔研究了这个问题，并于 1957 年发表了他们的结论。（后来他们又用计算机给出了更好的证明。）玩家的最佳策略我们已经介绍了，即在 9/11 的概率下决定要第三张牌，而在 2/11 的概率下决定不要第三张牌。庄家的最佳策略是按照 859∶1 429 的概率决定要或不要第三张牌，即在 3/8 的概率下决定要第三张牌，而在 5/8 的概率下决定不要第三张牌。弗兰克·道顿和罗格·霍德在 1972 年分析了多种赌场游戏，他们甚至给出了庄家如何确定具体策略的方法：当面对局面（1）时，根据自己手中的牌的花色做决定。 212

如果玩家选择最佳策略，则不论庄家如何选择，庄家的平均利润率都是 1.28%；同样，如果庄家选择最佳策略，则不论玩家如何选择，庄家的平均利润率都是 1.28%。可是，如果玩家偏离了最佳策略，庄家就有机可乘。例如以下三种情况：

○ 玩家在拿到 5 点时总是决定“不要”，则庄家在局面（1）下决定“不要”，平均利润率达到 1.54%；

○ 玩家在拿到 5 点时总是决定“要”，则庄家在局面（1）下决定“要”，平均利润率达到 1.37%；

○ 玩家在拿到 5 点时在 1/2 的概率下决定“要”，则庄家的平均利润率也可以达到 1.37%；

你可能觉得最佳策略产生的影响微乎其微，不过它毕竟可以提高你的收益。如果你不会在实战中应用最佳策略，你最好指望庄家也不会。也许你会遇到一个愚蠢的庄家，这家伙在面对局面（1）时总是决定“要”或总是决定“不要”。这时你可以利用他的愚蠢。如果他在局面（1）下总是决定“要”，你的

最佳策略是拿到5点时总是决定“不要”；如果他在局面（1）下总是决定“不要”，你的最佳策略是拿到5点时总是决定“要”。你在这个游戏中的平均收益肯定是负数，但最佳策略可以改善你的境遇。一点微弱的改善总不是坏事。

其他游戏

在赌场中，有1/6的赌注押在“二十一点”上，有1%的赌注押在掷骰子上。我们在第五章分析过掷骰子的游戏。二十一点吸引了许多数学家的兴趣，很多人在计算机的支持下研究这种游戏。1962年，爱德华·姚出版了《战胜庄家》一书。通常认为，如果玩家始终坚持冷静地分析和精确地计算，他可以在游戏中占据微弱的优势。马丁·密尔曼在1983年指出，如果游戏用4副牌进行，玩家的利润率可达1.35%；如果游戏用6副牌进行，玩家的利润率可达0.91%。不过玩家的最佳策略是非常复杂的，玩家需要根据自己手中的牌和庄家露出的牌作出各种决定，最佳策略由一个庞大的矩阵构成。而且，即使玩家可以掌握这种技术，他的优势也很微弱。他在游戏过程中失利的风险也是存在的。

213

如果你想掌握在这个游戏中制胜的绝技，请研究有关专著。你不仅需要仔细阅读，而且应当逐一证明作者给出的结论，以确保真正领悟神髓。我非常钦佩A·P·希尔伯特，这个古怪的天才曾经是无党派议员。此人曾以“农业问题”为主题发表竞选演讲，其纲领非常简洁：“我对农业一无所知。”这本书的目的是帮助你理解概率，而不是教你如何通过赌博致富。关于‘二十一’点，我的纲领也是明确而简洁的：“我对‘二

十一’点一无所知。”

不过我可以向你介绍制胜策略的概要。核心原则是通过对牌盒中剩下的牌的分析，判断形势是否对你有利。如果形势对你有利，则下最大赌注；如果形势对庄家有利，则不下注或下最小赌注。在这个游戏中，你应当要么下重注，要么不下注（或下最小赌注），绝不居中。

赌场对庄家的指示是，如果庄家手中为 17 点或更高，则不再要牌；如果庄家手中不超过 16 点，则再要一张牌。你希望庄家“涨死”。如果庄家手中的点数刚好略低于 17，而牌盒中剩下的小牌比较少，大牌比较多时，庄家涨死的概率较大。密尔顿介绍的方法是通过计算一个参数 T 来判断庄家涨死的概率。每一次洗牌之后开始计算，最初规定 T 为 0。赌桌上每出现一张 2、3、4、5 或 6，则把 T 加上 1；每出现一张 A、K、Q、J 或 10，则把 T 减去 1；对 7、8、9 不予考虑。最后，把 T 除以牌盒中剩下的牌的张数，并乘以 52。T 越大，则形势对你越有利。

当 T 为负数时，不要下注；当 T 为小于 2 的整数时，下最小赌注；当 T 不小于 2 时，下最大赌注。

这些结论是如何得到的？当你手中有两张 9 而庄家露出的牌为 7 时你是否应当再要一张牌？这些问题本书无法回答，请参阅专门研究“二十一”点的著作。我已经说过了，我对“二十一”点一无所知。 214

练习题

1. 假设你最初有 20 英镑赌本，你的目的是在轮盘赌中把

钱变成 100 英镑。你的方法是每次都押红色数字，使用大胆的下注方式。你成功的概率是多少？

假设你换一种赌法。首先用 16 英镑押六个数字（赔率为 1 赔 5），如果失利则用剩下的 4 英镑押数字 0。利用这种赌法，你成功的概率是多少？

你是否可以设计一种更好的赌法？

2. 你有 100 英镑赌本，每次用 1 英镑押一对数字 {21, 24}。连续玩 1 个小时，大约 90 轮。平均而言，结果如何？在第 89 轮结束时，你赢钱的概率是多少？在第 91 轮结束时呢？

3. 在酒鬼回家问题中，如果街长为 100 步，酒馆距离家 20 步，酒鬼安全到家的概率是多少？平均他将在街上走多少步？把他的步长变成原来的一半，结论如何？

4. 在赌场扑克中，以下三种情况的结果如何：

a) 你持 {Q, Q, K, 10, 4}，并决定增加赌注。庄家持 {Q, Q, J, 10, 9}；

b) 你持三张 K，并决定增加赌注。庄家持三带二；

215

c) 你持三带二，并决定增加赌注。庄家持三张 K。

第十一章　赌马，赛马彩票与差额赌博

赌　马

赌马业伴随着赛马业一同兴盛，赛马业的收入70%来自于赌博。针对赛马的赌博行为兴起于约克郡，历史可以追溯到16世纪。有据可查的最早的职业赛马者是哈里·奥当，他在1794年开展了赛马业务。1853年，赌马业被取缔，只有直接参加赛马的人才有权赌马。而且赌注必须使用银行票据，下注行为必须通过书信、电话或电报完成。很明显，这些规定把参与者限定在富有阶层中。当然，这些参与者可以派人到大街上搜罗想赌马的人，接受他们的现金赌注，替他们下注。1960年，赌博业合法化，地下赌博市场迅速衰退，而同时贿赂警察、诈骗保险金、团伙暴力等犯罪行为迅速兴盛。

最初政府不对赌博行为征税，赌博公司公开营业，不必担心来自犯罪团伙的不正当竞争。但随后政府开始征收赌博税，

并且税率越来越高。1978 年，皇家赌博调查团指出，过高的赌博税将迫使赌博业再次转入地下状态，破坏政府规范赌博业的企图。此后，政府降低了赌博税。现在的规定是，赌博中的获胜者缴纳赌博收入的 9%。这个比例比 1981 ~ 1996 年的比例低很多，但仍然有人认为这个税率超过了市场承受力。为了鼓励赛马业的发展，政府规定赌马免税。

现在英国有 8 000 多家赌博公司，其中的 40% 规模较小，而其他赌博公司则是属于四家大型赌博公司的子公司，最大的两家赌博公司是莱德布洛克公司和威廉·希尔公司。除赌马以外，赛马彩票和足球也是赌博公司的主要业务。大英帝国彩票兴起以后，赌博公司的业务受到很大的冲击，几家大型赌博公司联合起来开展了与大英帝国彩票类似的业务。赌博公司的业务包罗万象，你甚至可以针对你的孩子所在的运动队的成绩下注。但是如果你想赌海湾战争爆发的日期，没有哪家赌博公司肯接受你的下注。

216

赌博公司针对每个业务规定不同的赔率。赔率通常表示为 5:1，4:6，15:8 之类的形式。你肯定已经知道这些比例的含义，不过为避免混乱，我还是明确一下：“:”后面的数字表示顾客的赌注额，“:”前面的数字表示顾客在获胜的情况下赢得的奖金。例如，某匹马赔率为 4:6，意思是如果你下 6 英镑赌注，而这匹马果然获胜，则你的奖金额为 4 英镑。

赌博公司根据自己的判断确定赔率。假设某一场赛马中只有两匹马（a 和 b）参赛，赌博公司认为两匹马的胜率分别为 70% 和 30%，但是赌博公司绝不能规定两匹马的赔率为 3:7 和 7:3。赌博公司必须使赔率对自己有利，以确保利润。比如说，赌博公司可以规定 a 的赔率为 1:3，而 b 的赔率为 2:1。如果某个顾客对两匹马的胜率的估计与赌博公司的估计相同，他会

认为这两匹马都不值得下注。然而很可能有另一个顾客认为a的胜率高于75%，此人会认为1:3的赔率是对他有利的。就长期而言，如果他的判断是正确的，他会有所收益。另外，可能有人认为b的胜率高于1/3，这个人会乐于接受b的赔率。

赔率一定被设计成对赌博公司有利的形式。我们用“收益率”这个概念表示赔率对赌博公司有利的程度。赌场的收益率越高，顾客在赌博中越难获利。我们用赛马中的例子说明如何计算赌博公司的收益率。假定一场赛马中有n匹马参赛。如果第一匹马的赔率为x:y，则用y/（x+y）表示这匹马的收益率。分别计算出n匹马的收益率，并把这n个数加在一起，得到的和一定大于1，超出1的部分就是赌博公司在这场赛马中的收益率。例如，在我们刚才讨论的例子中，马a的赔率为1:3，于是其收益率为3/（1+3）=0.75；而马b的赔率为2:1，于是其收益率为1/（2+1）=0.3333。两个值的和为0.75+0.333 3=1.083 3，所以，赌博公司在这场赛马中的收益率为 217 8.33%。（事实上赌博公司的利润率比收益率略低一点。假设有75英镑押马a获胜，而33.33英镑押马b获胜，则无论哪匹马获胜赌博公司的支出额都是100英镑，而赌博公司的收入是108.33英镑，所以赌博公司的利润率为8.33/108.33=7.7%，略低于8.33%。不过收益率可以帮助我们很好地估计利润率。）在一场足球赛中，如果主队获胜、客队获胜及平局这三种结果的赔率分别为5:6，11:4，9:4，则赌博公司的收益率为6/11+4/15+4/13-1=0.119 8，接近12%。在大英帝国马赛中，参赛的马非常多，比赛结果的不确定性相当大，赌博公司的收益率也非常高。赌博公司在选择赔率时需要均衡考虑：太高的收益率会挫伤顾客的参与热情，而太低的收益率则会使赌博公司面临亏本的风险。赛马的收益率经常以赛马彩票的收益

率为参照，后者的收益率通常是19%。

赌场在一个业务中的收益率不可能是负数。假设有一场足球赛，赌博公司规定：主队获胜、平局和客队获胜这三种结果对应的赔率分别是6:5、11:4以及11:4，此时赌场的收益率是负数。一个顾客可以用75英镑押主队获胜，用44英镑押平局，另外用44英镑押客队获胜。这样，无论比赛的实际结果如何，这个顾客都可以保证有2英镑的进账。如果这个顾客把赌注额扩大1 000倍，这就是一笔不容忽视的财富。赌博公司决不允许这样的局面出现。当然，有时候赌博公司会把针对某一个事件的收益率设计为负数，但背后的圈套是通过其他业务挽回损失，并狠狠地赚一把。

有时候，在一场比赛中某一匹马获胜的可能性很大。此时赌博公司将设计一种赔率抵消这匹马对赌博结果产生的影响，在这种情况下会出现赔率为负数的情况。例如，在一场比赛中赌博公司的收益率为12%，而这场比赛中马A获胜的可能性很大，所以赌博公司把马A获胜的赔率设计成对本公司非常有利的形式，规定这匹马获胜的赔率为6:4，于是在这匹马上赌博公司的收益率高达4/（4+6）=40%。但是在整体上赌博公司的收益率仅为12%，所以在其他马身上赌博公司的收益率就是负数，即-28%。在这种情况下，如果你小心地把赌注押在马A以外的其他马身上，你可能赚一笔钱。总之，这种设计的目的是抵消马A的高胜率对赌博结果的影响。如果你偏偏愿意在马A上下重注（明知道这种情况对赌博公司有利），那也由你。218

正如在轮盘赌中赌场老板总是千方百计地使每个数字出现的概率相等，赌博公司也精心地设计每一匹马的赔率，使得无论顾客在哪一匹马上下注情况总是对赌博公司有利。下面这个

故事可以帮助你理解这个问题：

1996年9月28日，一个名叫弗朗凯·底特律的骑师在埃斯科特赛马会上夺得全部的七块金牌。在此之前，高顿·理查德在1933年曾连续获胜12次，但有一些比赛高顿没有参加；另外，一个叫威利·卡尔松的骑师曾在1990年的纽科斯特赛马会上夺得全部七块金牌中的六块。然而，弗朗凯·底特律的辉煌成就是史无前例的。对于赌博公司而言，这是十足的灾难，据估计，赌博公司因此损失了4 000万英镑。赌博公司的发言人声称："一名骑师夺得五块金牌就足以使赌博公司蒙受重大损失；夺得六块金牌就意味着灾难；如果夺得七块金牌，我们的老板就只能跳楼了。"几天之后，赌博公司终于从噩梦中缓过神来，公开赞扬弗朗凯·底特律的胜利是"赛马史上最辉煌的成就"。4 000万英镑的奖金全部兑现，各大媒体争相报道这一事件，并且向赌博公司免费提供宣传版面。弗朗凯·底特律事件剧烈地刺激了国民参与赌马的兴趣。

许多因素凑在一起造就了赌博公司的灾难。第一个因素是骑师本人。弗朗凯·底特律是一个罕见的天才，深受广大赌马者的爱戴。弗朗凯·底特律的每一次出场都吸引了巨额赌注，无论他骑哪匹马。许多与弗朗凯·底特律同场参赛者也愿意在弗朗凯·底特律身上下注。在历史上，还没有哪个骑师像他那样鼓舞下注者的热情；第二个因素是比赛规格。埃斯科特赛马会是英国最重要的赛事之一，这种档次的比赛总是吸引大量的参赌者；第三个因素是电视台的影响。BBC公司对这场赛事进行现场直播，这也是直接吸引赌马者的重要原因。

在比赛之前，赌博公司的专家估计底特律可能赢得七场比赛中的两场，最多不会超过三场。在七场比赛中，只有一场底特律被赌博公司当做夺标的热门，并设计了对赌博公司有利的

赔率，而在另外三场比赛中，赌博公司对底特律规定的赔率达到 10:1 甚至更高。赌马者既可以在场外下注，也可以在场内下注。如果赌马者下注的时间较早，则只能按照赌博公司已经规定的赔率下注；可是在临近比赛开始时，赌马者要按照即时赔率下注。某一场比赛的即时赔率是这一场比赛开始之前，根据比赛进行的情况临时决定的。在比赛的当天，现场有六七个赌马经纪人负责处理客户的场内下注，赌博公司根据这几个经纪人的收支情况决定即时赔率。赌博公司在全国各地设立了很多代办处接受投注，绝大多数投注行为是在场外进行的，不对即时赔率产生直接影响。各个代办处与场内保持密切的通讯联络，根据自己的情况要求公司调整即时赔率。

219

赌博公司事先规定的赔率反映了他们对比赛结果的预测。如果他们发现某一匹马吸引了非常多的赌注，他们会降低这匹马的即时赔率，以抵消自己承担的风险。当然，他们同时会提高其他马的即时赔率，以吸引客户继续投注。这种行动需要协调数百家场外代办处，难度可想而知。赌博公司经常发现，他们根本无法保证无论比赛出现什么样的结果都可以处于盈利状态。典型的情况是这样：某一场比赛有 20 匹马参加，赌博公司可以保证如果其中的 16 匹马中的 1 匹获胜，则公司盈利；如果不幸另外 4 匹马中的 1 匹获胜，则公司亏本。在极端情况下，赌博公司只能把希望寄托在 1 匹马身上，如果这匹马获胜则公司盈利，否则认赔。

在第一场比赛中，底特律乘一匹名为“华尔街”的坐骑出场。最初这匹马的赔率在 5:2 到 3:1 之间变化，但在比赛开始之前赔率变成 2:1。最终，华尔街超出第二名半个马身夺得锦标。这种现象不足为奇。第二场比赛底特律的坐骑名为“胆小鬼”，一共有 12 匹马参加这场角逐。当天早晨，胆小鬼的赔率

是 10:1，在开赛之前，赔率增加到 12:1。最后胆小鬼比第二名领先了一个马头。

第三场比赛在 3 点 20 分进行。这场比赛只有七匹马参加，但决定赌博公司命运的恰好是这场比赛。底特律将骑一匹名为“敬意”的马参赛。由于底特律已经赢得了前两场比赛，赌马者对底特律的信心大增，在敬意身上疯狂下注。在这场比赛中，赌博公司的整体收益率为 12%，敬意的赔率是 10:3。在全部七匹马中，敬意的赔率与另一匹马并列第二低。根据赌博公司事先的计算，如果敬意获胜公司还是可以盈利。但是由于底特律在前两场中获胜，大批赌注增加在敬意身上，局面已彻底改变。最终敬意超出第二名一又四分之一个马身。敬意获胜之后，赌博公司突然发现自己已面临困境。数以百万计的电视观众目睹了底特律的胜利，底特律变成一个不可战胜的英雄。 220

3 点 55 分，底特律乘“受勋英雄”第四次出场。当天早晨，受勋英雄的赔率是 14:1，但是在上一场比赛结束后的几分钟之内，赔率迅速跌到 8:1。这场比赛有 26 匹马参加，从赌博公司的角度看，这个赔率已经够保险的了。在比赛开始之前，赔率又跌到7:1。最后受勋英雄以三个半马身的优势胜出，同时赌博公司的亏损额以火箭的速度飙升。对于那些已经受理的投注赌博公司无能为力，目前他们惟一能做的事就是狠狠地降低底特律在以后几场比赛中的赔率。

赌博公司寄希望于场内的经纪人。各大场外代办处纷纷与总部联络，要求调整赔率。情况非常危急，赌博公司不得不动用储备资金向客户支付奖金，但是储备资金不能立刻到位，因为亏损的窟窿太大了，如果立即动用储备资金，储备资金会迅速耗尽。底特律乘“宿命”第五次出场时，宿命的赔率已从原来的 10:3 降到 7:4，但仍然吸引了巨额赌注。底特律又以微弱

的优势获胜。

场内经纪人的努力还是有成效的。赌博公司亏本的命运无法改变，但是由于及时调整赔率，他们的损失额降低了不少。另外，场内经纪人自己也在底特律身上下了重注，赢得了一笔钱。

底特律连续获胜五次，引起了电视台的关注。BBC 本来计划在五点钟转播足球比赛，现在放弃了原定计划，继续直播赛马实况。赛马场成为全国瞩目的焦点。大批赌注再次投向底特律，赌博公司不得不再次降低赔率。在开赛前两分钟，底特律的第六匹坐骑的赔率为 13:8，而最终的赔率为 5:4。底特律第六度获胜，领先于第二名四分之三个马身。

最后一场比赛在五点半举行。底特律乘“富士山绝顶”出221 场。底特律在前一年曾赢得这个比赛的冠军，当时他的坐骑也是富士山绝顶。富士山绝顶以往的战绩不理想，曾参赛五次，有两次没有进入前三名。这场比赛有 18 匹马参赛，富士山绝顶最初的赔率为 12:1。在第四场比赛结束之后，赔率已跌到 7:1；在第五场比赛结束之后，赔率又跌到 4:1；当底特律第六次获胜时，赔率竟然跌到 6:4。场内经纪人在确定即时赔率时，并不考虑本公司在以前的业务中的盈亏情况，只是根据自己对比赛的预测以及顾客的投注情况确定策略。当赔率达到 7:4时，场内经纪人决定把赔率提高一点，吸引更多的赌注。最终赔率为 2:1。他们这样做是因为他们不相信底特律会赢得全部七场比赛，在历史上还没有人实现如此辉煌的业绩。（正如在轮盘赌中连续出现 10 次红色数字以后，庄家不相信下一次也出现红色数字。）但是他们想错了。底特律完成了辉煌的七连胜。

观众在欢呼，赌博公司却陷入无边的烦恼。他们惟一的安

慰就是由于及时调整即时赔率使得损失减少了一部分。有些经纪人在庆幸自己的聪明：他们通过在底特律身上下注自己赢了一些钱。有些顾客下注赌底特律在全部七场比赛中获胜。赌博公司事先规定的赔率为23.5万:1，后来即时赔率调整为2.5万:1。这下子赌博公司输惨了。

有一个年轻人用64英镑赌底特律赢得全部七场比赛，结果一下子赢回55万英镑！这个乐观的家伙在第二天中故伎重施，又用64英镑赌底特律赢得全部七场比赛，不过这次底特律只赢了一场。

这种罕见的情况使得赌博公司损失惨重，但是在200年的赛马史上发生一次这种情况也不奇怪。其实赌博公司的角色很像保险公司：如果保险公司接受了对于地震、飓风等天灾的投保，而投保事件确实发生（虽然其概率极低），保险公司只有咬着牙承担损失。在这个例子中，赌博公司的困境也是如此。他们已尽了最大努力，但依然无法逃避灾难。

222

表11.1：第三场比赛的盈亏统计

参赛者	投注额（英镑）	即时赔率	理想的即时赔率
A	10 500	9:4	4:1
B	21 158	10:3	6:4
C	9 080	10:3	9:2
D	6 334	11:2	7:1
E	1 998	10:1	25:1
F	6 430	14:1	7:1
G	2 944	25:1	18:1
合计	58 444	1.119	1.123

下面我们从赌博公司的立场出发分析得失情况。以第三场比赛为例，如表 11.1 所示。第一列表示参赛的七匹马，第二列表示每匹马吸引的赌注，赌注总额为 58 444 英镑。第三列表示每匹马对应的即时赔率。无论哪匹马获胜，我们都可以很容易地算出赌博公司的盈亏状况。例如，如果马 A 获胜，赌博公司需要支付$\frac{9+4}{4}\times 10\ 500 = 34\ 124$英镑，其中包括退还给获胜的客户的赌注。赌博公司接受的赌注总额为 58 444 英镑，所以赌博公司盈利 58 444 − 34 125 = 24 319 英镑。如果获胜者是 C、D 或 E，赌博公司盈利；如果获胜者是 B、F 或 G，赌博公司亏损。对赌博公司而言，最糟的结果是马 F 获胜（亏损 3.8 万英镑），最好的结果是马 E 获胜（盈利 36 500 英镑）。

如果赌博公司可以在拿到所有统计数据之后自由决定赔率，他们最好把赔率设计成表中最后一列的形式。在此条件下，无论比赛结果如何，公司都能盈利，最糟的结果是 G 获胜（盈利 3 400 英镑），最好的结果是 C 获胜（盈利 8 350 英镑）。

C 就是底特律的坐骑——敬意。单从表 11.1 看，敬意获胜对赌博公司有利。但实际上，有很多客户的赌注同时赌两场或更多场比赛的结果，这些赌注都与这一场比赛的结果相关。从总体看，赌博公司的亏损远远高于收益。如果把所有赌注考虑在内，再加上通讯效率的限制，为赌博公司设计一个只盈不赔的理想赔率体系是不可能的。

223

我们可以用同样的方法分析当天的全部七场比赛。忽略那些同时赌多场比赛结果的赌注，只考虑直接押在每一匹马身上的赌注，则在全部 93 匹赛马中，有 75 匹马的胜利意味着赌博公司盈利，而另外 18 匹马的胜利意味着客户盈利。在每场比

赛中，至少有一匹马可以使赌博公司亏本。

赌马者最关心的问题是对什么样的马下注最有利。比如说，有一匹马是夺标热门，赔率为5:4，而另一匹马获胜的可能性较低，但赔率高达66:1，客户应当选择哪匹马下注？1978年的皇家委员会讨论过这个问题的细节，其后一些职业的统计学专家和经济学家深入研究了类似的问题。最后的结论是：平均而言，选择赔率较低的马更好。

这个问题没有一个简单的答案。皇家委员会研究了25年积累的统计数据，得出一个结论：如果一个客户坚持对每一匹赔率为4:6或更低的马下注，则总体而言此人可望获得微弱收益（不考虑赌博税）；如果一个客户坚持对每一匹赔率在9:1到20:1之间的马下注，则平均而言此人将损失赌本的35%；如果一个客户坚持对每一匹赔率高于20:1或更低的马下注，则平均而言此人将损失赌本的70%。根据1973年的统计数据，在4 000匹赔率为33:1或50:1的马中，只有23匹获得胜利。在这种胜率下，即使赌博公司的赔率为100:1，赌博公司仍然可以保证客观的利润。根据最近的统计数据，如果客户对赔率较低的马下注，基本可以保持收支平衡，但如果对赔率达到10:1或更高的马下注，平均将损失30%以上。

有几种理论可以解释这一现象，不过这几种理论是互相抵触的。一种理论的出发点是把赌马与股票交易相类比：在所有客户中有一部分人是庄家，他们可以相当精确地预见哪一匹马将获胜，这些人的选择造就了一部分马的低赔率，所以低赔率的马胜率极高。另一种理论则分析客户与赌博公司在对抗中的策略选择：双方都把赌马当做一种赌场游戏，从整体上说游戏规则对赌博公司有利。赌博公司为了赚钱，必须吸引客户下注，所以在这个游戏中赌博公司不能把自己的优势设计得太

大。客户当然想在游戏中赢钱，但他们的主要兴趣在于用一笔小赌注赢得一笔大奖金。所以即使客户明知某一匹马获胜的概率不足1%，他们也乐于接受33:1的赔率，而赌博公司当然不会提供比33:1更慷慨的赔率。从另一个角度看，如果某一匹马在比赛中明显占优，胜率超过三分之二，1:2的赔率就是对客户有利的。而赌博公司为了吸引客户的兴趣，还要把赔率提高一点，比如说4:7。这只是非常粗略的描述，实际情况当然复杂得多。赛马和掷骰子、抛硬币不同，精确地判断某匹马的胜率是一门复杂的艺术。不过赌博公司已经掌握了一个简单的真理：他们的主要收入通常不是来自于那些出类拔萃的骑师。

在足球业务中，赌博公司对局面的控制更加有力。一个典型的例子：在1997年9月的一个周末，共有66场足球赛，在每场比赛中，赌博公司的收益率都在11%到13%之间。投注行为必须在这一周开始时完成。如果在投注之后的几天之内发生了某些意外变故，比如队员转会、受伤等等，这些因素都会改变客户最初对比赛结果的预计。这些不确定的因素是客户必须承担的风险。在足球业务中赌博公司设计了一些保护性措施。例如，如果一个客户想赌某一场比赛的结果是主队获胜，则根据赌博公司的规定，这个客户必须同时预测至少五场比赛的结果，只有在这五个预测都正确的前提下，客户才能拿到奖金。这样，赌博公司的利润同时来自于五场比赛，如果在每场比赛中赌博公司的收益率为12%，则在五场比赛中收益率就高达76%。如果一个客户想赌某一场比赛的结果是客队获胜，则这个客户只需要同时预测三场比赛的结果，但这样仍然是赌博公司的收益率高达40%。在某种情况下赌博公司会放宽这种限制，例如，如果一场比赛将由电视台现场直播，则允许客户单独对这一场比赛的结果下注，或同时对两场比赛下注（其

中一场是被直播的比赛）。

大英帝国彩票兴起以后，赌博公司感到了竞争的压力。现在他们也接受对一场比赛的投注。政府规定的赌博税高达9%，赌博公司必须向政府缴纳全部利润的40%。为了生存，赌博公司千方百计地增加对客户的吸引力。

法律从未禁止赌博公司经营国外彩票。爱尔兰足球彩票在英国已发行多年，每周开奖两次。几家赌博公司甚至联合起来发展了与大英帝国彩票形式相仿的业务。他们从法国买回同样的开奖机器，并且组织了中立机关监视彩票的公正性。

这种彩票的形式花哨而多样，但实质与大英帝国彩票相同。客户在49个号码中选择号码，可以选1～5个。如果客户选择的号码全部出现，则赌博公司支付奖金，赔率取决于客户选择了几个号码。很容易算出客户中奖的概率和投资回报率，细节如表11.2所示。

如果你选3个号码，赌博公司的赔率是511:1。假设你选择的是12、31和41（其实具体选择哪个号码无关紧要），12在结果中出现的概率是6/49，在12出现的前提下31也出现的概率是5/48，在12和31都出现的前提下41也出现的概率是4/47，所以你中奖的概率为$\frac{6}{49}\times\frac{5}{48}\times\frac{4}{47}=\frac{5}{4\,606}$。这个值低于1/920。由于赔率为512:1，所以你的投资回报率为$512\times\frac{5}{4606}=0.555\,79\cdots\cdots$。

表 11.2

客户选择的号码个数	中奖概率	赔率	投资回报率（%）
1	1/8.71	11:2	79.6
2	1/78.4	48:1	62.5
3	1/921.2	511:1	55.6
4	1/14125	656 0:1	46.4
5	1/317 814	999 99:1	31.8

根据表 11.1，你追逐的赔率越高，你的损失越大。在赔率中已经扣除了赌博税，所以如果你用 1 英镑赢回了 512 英镑，你不必再交税。另外，赌博公司规定了奖金额的上限，每张彩票享有的奖金不得超过 25 万英镑。这是一个保护性的规定。

这个规定可以帮助我们理解为什么赌博公司不允许同时选 6 个号码。在第二章中我们已经得出结论，同时选 6 个号码而中奖的概率为 1/14 000 000。按照表 11.1 中的规律外推，这种彩票的赔率应为 3 000 000:1，也就是说，如果客户的投注额高于 8 便士，奖金额就会超出预先规定的上限。一次偶然事件的发生就可能导致赌博公司亏本，赌博公司的宗旨是远离风险。

你可以利用赌博公司赚钱吗？

许多赌马者相信自己可以通过赌马发家致富（正如陀思妥耶夫斯基相信自己可以在轮盘赌中制胜）。通过赌马赚钱需要两个前提：丰富的知识（精确地判断各种局面对自己的有利程

度）和坚强的意志（控制自己冷静地下注）。即使你已经具备了这两个条件，你还需要解决一个问题：如果你已确知赌某匹马对你有利，你应当投入多少赌注？

约翰·L·克里在他的著作中回答了这个问题。粗看起来，既然局面对你有利，你应当在这个赌博中投入自己的全部资金，以获得最大回报。但从长远看，这种策略一定是失败的。因为只要一次不利的局面出现就会使你血本无归，而不论不利的局面出现的概率多么微小，长期而言不利的局面迟早会出现。最佳策略一定要比孤注一掷谨慎。

克里提供的最佳策略是，让每次投入的赌注在现有的全部资金中占有固定的比例。如果赌博确实对你有利，应用这种策略可以使你的资金呈几何级数增长，就像放高利贷一样。当然，这个比例的选择是有技巧的。如果比例太低，那么你的资金增长速度会非常缓慢；如果比例太高，你就要面对亏本的风险。一个适当的比例意味着在速度和风险之间取得完美的均衡，以确保利益最大化。克里的结论是：这个比例应当等于你在赌博中的优势的大小。例如，一匹马获胜的概率为51%，失败的概率为49%，赔率为1:1，则你的优势为2%，你应当拿出所有资金的2%赌这匹马获胜。如果这匹马获胜的概率为55%，则你的优势为10%，此时你应当拿出所有资金的10%下注。长期而言，这种策略的收益最大。

227

你必须耐心地赌许多轮。在这个过程中，你可能遇到多次挫折。如果一匹马的胜率为51%，而你遵循最佳策略每次下注2%，则你的资金平均每轮增长0.02%；如果一匹马的胜率为55%，而你每次下注10%，则你的资金平均每轮增长0.5%。你当然可以提高下注的比例，这样资金增长的速度会提高，但同时你破产的概率也增加了。

当赔率不是 1:1 时，你计算优势的算法要略作调整。比如说，某一匹马的赔率为 3:1，而这匹马获胜的概率高于 25%，则赌这匹马获胜显然对你有利，但是你的优势是多少呢？计算方法很简单。假定这匹马获胜的概率为 28%，则你获胜的概率为 28%，失败的概率为 72%，你的优势是 28% - 72%/3 = 4%。所以每次你应当投入全部资金的 4%。

另外两个例子：如果你获胜的概率为 34%，而赔率为 2:1，则你的优势为 34% - （1 - 34%）/2 = 1%，即每次你应投入全部资金的 1%；如果你获胜的概率为 70%，而赔率为 1:2，则你的优势为 70% - （1 - 70%）2 = 10%，即你每次应投入全部资金的 10%。（其实在这两个例子中算法完全一样，赔率为 1:2 也就是 1/2:1，而乘以 2 也就是除以 1/2。）

应用克里的策略时需要注意两种特例。第一种特例是你在实际应用这一策略时可能遇到小麻烦。比如说，你的赌本是 1 英镑，而根据克里的策略每次你要投入 10% 的资金。第一次你投资 10 便士，如果投资失败，你还剩下 90 便士；第二次你的投资额应当是 8.1 便士，但这是不可能的，你到哪去找 0.1 便士呢？在你的赌本数额非常大时，你也可能遇到这种麻烦。克里对这种情况的建议是把你的投资额减少一点点。如果投资额的计算结果是 8.1 便士，你就投资 8 便士；如果投资额的计算结果是 8.9 便士，你就投资 8 便士；当你的全部资金不足 10 便士时，你的投资额的计算结果低于 1 便士，你的投资额就是 0，此时你可以认为你已经破产。即使在对客户有利的赌局中，客户血本无归的情况也时有发生。

第二种特例是为了增加保险系数。克里的策略可以保证投资回报率最高，但是有些客户更加关心如何降低破产的风险。他们的方法是每次投入的赌注都比克里建议的略小一点。这样

做的结果是使得资金增长的速度降低，但风险确实得到了进一步的控制。在上一种特例中，克里建议把 8.1 便士和 8.9 便士都调整为 8 便士，就是基于这种考虑。当然你也可以使每次投入的赌注都比克里建议的略大一点，这样资金增长的速度将增加，代价是破产的概率同时上升。

面对一场具体的比赛时，有时你会发现对你有利的马不只一匹。比如说，有两匹马获胜的概率都是 20%，而赔率分别为 4:1 和 5:1，显然，对这两匹马下注都是有利可图的。这种情况也在克里的讨论之列。他的结论是，你应当同时在这两匹马身上下注。克里分析了在一般情况下如何下注才能使资金增长的速度最大化，但是结论出乎很多赌徒的意料：为了使资金增长的速度达到最大，有时候你应当在那些对你不利的马身上下注。

你觉得这个结论难以接受吗？考虑这个例子：一场比赛有 5 匹马（A，B，C，D 和 E）参加，赔率分别为 13:8，9:4，9:2，5:1 和 20:1，（在这场比赛中，赌博公司的收益率为 8%。）而你估计这 5 匹马的胜率分别为 40%、30%、20%、8% 和 2%。很明显，A 和 C 的赔率对你有利，而 B、D 和 E 的赔率对你不利。克里认为，最佳的下注方式是同时在 A、B 和 C 上下注。

克里用到的数学工具比较复杂，我不想在这里讨论其中的细节，有兴趣的读者可以钻研附录 5。克里提供的策略是每次拿出全部资金的 23% 下注：10.6% 押马 A，6.25% 押马 B，6% 押马 C。这种策略可以保证资金增长率为 0.35%。如果比赛重复进行 N 次，而每次你都这样下注，则你最终的资金额为最初的（1.003 5）^{N} 倍。 229

假如这个例子中马 B 的赔率改为 2:1，而马 A 和 C 的赔率

不变，马 D 和 E 的赔率略作调整，最佳下注方案如何？克里的结论是只对马 A 和 C 下注。具体地说，每次拿出全部资金的 5.15%押马 A，同时拿出 3.37%押马 C。这样可以保证资金增长率为 0.21%。很容易理解为什么你的资金增长率降低了，因为有一匹优势较大的马赔率下调，所以你的整体优势下降。

这个例子说明精确计算的结果往往与直觉相悖。当马 B 的赔率为 9:4 时，你同时在马 A、B 和 C 上下注，这三匹马中的一匹获胜的概率高达 90%，只有在 D 或 E 获胜的情况下你才赔本，概率只有 10%，所以你可以大胆一些，一次投入全部资金的 23%；而当马 B 的赔率为 2:1 时，你只在马 A 和 C 上下注，你获利的概率只有 60%，所以你的投资应当更加谨慎，每次仅投入全部资金的 8.5%。

由于马 B 的胜率为 30%，所以当马 B 的赔率低于 7:3 时单独对马 B 下注是不利的。但是 9:4 与 7:3 非常接近，事实上，同时对马 A、B 和 C 下注反而使资金增长率达到最高。

赛马彩票

赛马彩票始于 1929 年。政府批准赛马彩票出于两个考虑：其一，提供一种与赌博公司竞争的经营形式；其二，让赛马业通过彩票获得资金支持。赛马彩票的奖金分配规则与足球彩票类似，在全部销售收入中取出比例固定的一部分作为发行机关的利润和经营成本，其他部分则作为奖金返还购买者。当一个人购买赛马彩票时，实际上他是在同其他的购买者对抗，因为发行机构最终支付的奖金份额是固定的。发行机构绝无亏本的风险，因为总奖金额必然低于总投注额。

230

下面是一个典型的例子：鲍勃想买下一场马赛的彩票，这场比赛有10匹马参加。在下注窗口的上方有一个电视屏幕，当前的彩票销售情况随时显示在屏幕上。通过屏幕显示的数据，可以估计出每种下注方式对应的赔率。在销售的早期，这些数据变化得很厉害。鲍勃是一个谨慎的人，他耐心地研究这些数据，在比赛开始前的最后几分钟，他确定了一种他认为最有利的下注方式。他用两英镑买"甲虫炸弹"，同时押这匹马夺冠和进入前三名。

在这一轮彩票销售结束之后，发行机关的统计结果是：共有5 000英镑押某匹马夺冠，3 900英镑押某匹马进入前三名。在第一笔赌注中，发行机关要扣除16%，在第二笔赌注中，发行机关要扣除24%，扣除的部分作为发行机关的经营成本和利润。在这两笔赌注中，剩下的份额分别是4 200英镑和2 964英镑。这些钱就是支付给购买者的奖金。最后甲虫炸弹夺冠。

鲍勃得到多少奖金取决于有多少人和他分享奖金。假设在这一轮中共有1 000英镑的赌注押甲虫炸弹夺冠，而对应的奖金额为4 200英镑，所以每1英镑的赌注享有4.20英镑的奖金。鲍勃投入了2英镑的赌注，所以他得到8.40英镑的奖金。如果在这一轮中只有100英镑的赌注押甲虫炸弹夺冠，则每1英镑的赌注享有的奖金额上升到42英镑，鲍勃收入84英镑。

那些押某匹马进入前三名的赌注用同样的方式处理。全部奖金额为2964英镑，但进入前三名的马有三匹，所以冠军、亚军和季军分别对应988英镑的奖金。假设共有500英镑的赌注押甲虫炸弹进入前三名，则在这些赌注中每1英镑赌注对应的奖金额为988/500，约等于1.9英镑。鲍勃投入的赌注为2英镑，所以他的收入是3.8英镑。如果只有100英镑的赌注押

甲虫炸弹进入前三名，则每 1 英镑赌注对应的奖金额上升到 9.8 英镑。押在亚军和季军身上的赌注的赔率可能比冠军高。

这种规则可能造就令人尴尬的结果。在这个例子中，如果押甲虫炸弹夺冠的赌注为 1 000 英镑，而押这匹马进入前三名的赌注只有 100 英镑，结果是押这匹马夺冠的回报率为 4.20:1，但押这匹马进入前三名的回报率却高达 9.80:1。这种事在美国发生过。那是 1973 年，一匹名为“秘书处”的马在一场比赛中轻松获胜。如果你用 2 美元押这匹马夺冠，你的奖金为 2.20 美元，但如果你用 2 美元押这匹马进入前三名，你的奖231金为 2.40 美元。有些聪明的购买者预见到这种局面，他们明知这匹马夺冠的可能性极高，但并不押这匹马夺冠，而是押这匹马进入前三名。当参赛的马有 5～7 匹时，购买者可以押某匹马进入前两名；当参赛的马有 8～15 匹时，购买者可以押某匹马进入前三名；当参赛的马有 16 匹或更多时，购买者可以押某匹马进入前四名。

如果你押某匹马夺冠，你的投资回报率为 84%；如果你押某匹马进入前几名，你的投资回报率为 76%。如果你按其他方式下注，你的投资回报率在 71% 到 74% 之间。如果一个购买者预测当天的前六场比赛的冠军，并且预测正确，则此人获得头奖。如果某一天没产生头奖，则头奖奖金自动累加到次日的头奖奖金额中。这个规定和大英帝国彩票中的规定相同。

如果赛马彩票提供的奖励力度为 1 英镑奖励 12 英镑，情况与赌博公司提供 11:1 的赔率完全相同。有些赌马者因此要求赌博公司按照赛马彩票的赔率确定赔率，赌博公司拒绝了这个要求。赌博公司的理由是，他们提供的即时赔率比赛马彩票的赔率更公平。不过皇家委员会不同意这个理由。皇家委员会指出，赌博公司的即时赔率不过是根据一些体育记者的观点制

定的，而赛马彩票的赔率是通过购买者的行为自动生成的。

差额赌博

差额赌博是一种起步较晚但发展迅速的赌博形式。这种赌博类似于期货交易。比较一下期货交易市场中的例子：有一批咖啡豆将在六个月内到货，期货市场内的价格为 1 500 ~ 1 600 英镑。1 500 ~ 1 600 英镑的含义是指卖出方必须以 1 500 英镑的价格出货，而买入方必须以 1 600 英镑的价格进货。买进（或卖出）的合同本身是可以在期货市场中买卖的。一个人在期货交易中的得失取决于合同价格与市场价格的比照。如果你按照比市场价格低的合同价格进货（或者按照比市场价格高的合同价格出货），你就赚了，反之你就亏了。差额赌博的原理与期货交易相同，差别在于在差额赌博中炒作的对象不是咖啡豆的市场价，而是体育竞赛的结果。

下面这个例子是赌澳大利亚板球队在与英国队对阵时第一局将得多少分。这个交易的期货价格为 340 ~ 360，客户可以选 232 择买进或卖出。如果一个客户 A 预测澳大利亚队将打得很糟，他可以卖出，卖出价为 340。假设他以 2 英镑/分的赌注卖出，而比赛的结果是澳大利亚队的得分确实低于 340 分，他就赚钱了，分数每低 1 分他可以得到 2 英镑。如果实际得分为 300 分，他的收入为 80 英镑。如果澳大利亚队的实际得分高于 340 分，他就亏本了，分数每高 1 分他就亏损 2 英镑。如果实际得分为 450 分，他亏损 220 英镑。相反，另外一个客户 B 预测澳大利亚队将打得很好，他可以买进，买进价为 360。假设他以 3 英镑/分的赌注买进，而比赛的结果是澳大利亚队的得分确

实高于360分，他就赚钱了，分数每高1分他可以得到3英镑。如果实际得分为450分，他的收入为270英镑。如果澳大利亚队的实际得分低于360分，他就亏本了，分数每低1分他就亏损3英镑。如果实际得分为300分，他亏损180英镑。细节见表11.3。

正如赌马活动的经营者精心设计赔率以对抗客户，差额赌博的经营者也精心设计期货价格以对抗参与者。在差额赌博中，经营者可以通过三种情况获得利润：

其一，实际结果恰好在期货价格的最高价和最低价之间。在刚才的例子中，如果实际得分高于340分而低于360分，为345分，则经营者两面收钱。选择卖出的客户的损失是“卖出量”的5倍，选择买进的客户的损失是“买进量”的15倍，这些钱进入经营者的腰包。（这种情况类似于在期货市场上实际价格高于卖出价而低于买进价，买方和卖方都受损。）

其二，客户的“买入量”与“卖出量”持平。如果买进的赌注总额与卖出的赌注总额都是1 000英镑，则不论最终的实际结果如何，经营者都可以保证盈利，利润额为买入价与卖出价的差价乘以1 000。这种情况类似于在一场赛马中如果每匹马吸引的赌注一样多，则赌博公司必然盈利。在赌马活动中，赌博公司可以通过变换赔率影响每匹马吸引的赌注，在差额赌博中经营者也可以通过调整期货价格影响买入量和卖出量。比如说，当经营者发现卖出量明显高于买进量时，他们可以下调期货价格，从而鼓励买进，抑制卖出。

其三，经营者比客户更准确地预测比赛的结果。例如，某一场比赛的期货价格为340~360，多数客户认为买进比较合算，买进量为卖出量的三倍。这时，经营者可以上调期货价格，比如说把期货价格调整为390~410，通过这种行为抑制买

进，鼓励卖出。但是经营者不一定采取这种策略。如果经营者认为客户的判断是错误的，他们就会坚持最初的价格。如果最终的事实验证了经营者的判断，他们就发财了。不过如果经营者的判断是错误的，他们将付出惨重的代价。在实际的期货交易中，这种灾难屡见不鲜，比如错误的判断导致巴林银行倒闭。

表 11.3：期货价格为 340 ~ 360，A 以 2 英镑/分的"量"卖出，B 以 3 英镑/分的"量"买进

实际分数	250	300	350	400	450
A 的收入	180	80	- 20	- 120	- 220
B 的收入	- 330	- 180	- 30	120	270

233

赌马活动的经营者在设计赔率时必须保证对赌马者的吸引力，同样，差额赌博的经营者也面对同样的问题，他们必须保证期货价格同时对卖方和买方保持吸引力。经营者不可能做到每次都赚钱，但妥善经营可以使他们赚多赔少，赚大赔小，这样从长期看经营者就可以确保可观的利润。确定期货价格的实际程序是什么样的呢？请看这个例子：

经营者要为某一支足球队在某场比赛中的第一次进球的时间设计期货价格，他们首先要利用统计学手段算出这个时间的期望值。根据预测，这支球队在这场比赛中的进球数将在 2 和 3 之间。假定随着比赛的进行，这支球队进球的概率均匀上升，于是第一次进球的时间与进球数成反比。整场比赛 90 分钟，如果经营者预测这支球队将进 2.6 个球，则第一次进球的时间为 90/2.6，约等于 34.6。我们有理由认为第一次进球可

以使以后的进球容易一些，所以可以把这个值放大一点，比如说 36，即第一次进球的时间的期望值为 36 分钟。如果经营者预测进球数大于 2.6，这个期望值要下调，反之则上调。

在期望值确定以后，就可以设计期货价格。最基本的原则在于，期望值应当出现在期货价格的范围内，即高于卖出价而低于买进价。从长期看，这种设计使得经营者同时从买进者和卖出者身上赚钱。如何规定买进价和卖出价之间的差价很有学问。单纯从经营者的利益出发，当然是差价越大越有利。差价越大，经营者承担的风险越小。另外，在买入量和卖出量持平的情况下，经营者的收益固定等于买入量（或卖出量）乘以差价，所以较高的差价意味着较高的利润。但是相反的因素必须考虑在内，即随着差价的提高，客户参与交易的兴趣下降，而经营者为了赚钱，必须吸引足够多的客户参与。我们可以粗略地认为，差价降低一半使交易量提高一倍。

通常，随着比赛的进行，期货价格将活跃地变动，而客户可以继续参与交易。第二次交易通常可以保护客户的利益。还是用板球赛的例子。最初的期货价格为 340~360，而在半场休息时，澳大利亚队已经得到 220 分。显然，球队的表现比经营者的预期好，于是经营者把期货价格调整为 420~440。客户 A 事先以 2 英镑/分的赌注买进，当时的买进价为 360。此时他可以以 2 英镑/分的赌注卖出，而卖出价为 420。这样，无论终场得分为多少，他都赚到了 120 英镑。事实上，他的交易活动已经结束了。我们来分析一下第二次交易的后果。如果终场得分为 260 分，而 A 没有做二次交易，则他要输 200 英镑，但是做第二次交易使得他最终赚了 120 英镑，所以第二次交易非常成功，收益额为 320 英镑。当然相反的情况也可能发生。如果终场得分为 500 分，而 A 没有做二次交易，则他将赢 280 英镑，

234

但是做第二次交易使得他只赚到 120 英镑，所以第二次交易失败，损失额为 160 英镑。但是无论终场得分为多少，第二次交易使得 A 的收益额固定为 120 英镑。假设实际的比赛以另一个形式进行。在半场休息时，澳大利亚队的得分只有 160 分，表现比经营者的预期差，于是期货价格下调为 230～245。客户 A 事先以 2 英镑/分的赌注买进，当时的买进价为 360。现在 A 发现自己处于不利的形势中，此时他可以以 2 英镑/分的赌注卖出，卖出价为 230。当然他赔本了，但是无论终场得分为多少，他的损失额是固定的，260 英镑。如果实际的终场得分为 180，则第二次交易使得 A 的损失减少了 100 英镑。最初买进（或卖出），其后又卖出（或买进），这种情况称为“反向操作”。反向操作可以使客户的得失与实际的比赛结果无关。

参与差额赌博有一个很明显的风险，即损失额是未知的。当客户要求买进（或卖出）时，他并不是以固定的赌注额下注。在板球赛的例子中，如果你以 2 英镑/分的赌注卖出，但澳大利亚队的实际得分超过 700，我可以设想你的心情。当然，如果你最初的选择是以 2 英镑/分的赌注买进，你的收入也很可观。正如报纸上经常刊登“XX 政府忠告市民，股票有风险”的广告，差额赌博的风险也是巨大的。有时候，经营者允许参与者设定输赢的上限，但是你自己应当清楚你的经济承受能力。

如果你想在差额赌博中对一场足球赛下注，下注的方式可以有许多种。最常见的下注方式是赌双方的总得分或某一个队的净胜球。在这类情况中，得分数（或进球数）在期货价格中以分数形式表示，精确到十分之一。例如，在一场比赛中，你想赌利物浦队的净胜球数，经营者最初提供的期货价格为 0.6～0.9。随着比赛的进行，期货价格不断变化。当利物浦队领

先两分时，期货价格变为2.7~3.0。你甚至可以对某一场比赛中每一个进球的队员的号码之和下注。还有一种更玄的下注方式：假定某一场比赛的进球数为a，角球数为b，比分为0:0的持续时间为c（以分钟为单位），有效进球数为d，黄牌数为e，红牌数为f，则可以确定一个值x：$x = 25a + 3b + c - 10d - 10e - 25f$。你可以对x的值下注，当经营者提供的期货价格为45~50时，你要决定买进或卖出。

那些只分胜负、不计得分的比赛也属于差额赌博的业务范围，具体的方法是为每一种结果规定一个得分，例如，规定全胜者得100分，终场获胜者得75分，半场领先得50分，等等。当一场比赛由两个队参赛时，经营者会精心设计每支队伍的期货价格，使得无论客户在哪支队伍上下注经营者都处于有利地位。例如，在1997年的橄榄球甲级联赛中，维京队对抗圣海伦队，双方的期货价格分别是92~94和81~83。如果你用1英镑的赌注买进维京队，维京队获胜将使你收入6英镑，维京队失败将使你损失19英镑，因此，经营者为维京队设计的赔率相当于6/19。类似地，经营者为圣海伦队设计的赔率相当于17/8。所以经营者的收益率为$19/25 + 8/25 - 1 = 2/25$，即8%。

经营者允许你对在某个下午举行的赛马会的全部比赛下注，下注方式多种多样。你可以赌某一匹马在一场比赛中领先另一匹马的距离（以马身的长度为单位），也可以赌某匹马的得分。计分标准由经营者事先规定，比如，规定冠军、亚军、季军的得分分别为25、10和5。你也可以对多场比赛的结果下注，计分标准比较复杂。你甚至可以赌在整个赛马会中所有夺冠的马的年龄总和。另外，经营者也设计了头奖。

假设一场比赛有20匹马参赛。经营者规定前四名的得分

分别为50、25、10和5，其他马的得分为0。每匹马的期货价格由它获胜的概率决定。具体情况可能是这样：夺标第一热门的期货价格为22～24，夺标第二热门的期货价格为15～17，随着获胜概率的下降期货价格递减，最不被看好的马的期货价格为0～1.5。经营者可以根据赌博公司对每匹马规定的赔率确定每匹马的期货价格，程序如下：

1. 利用每匹马的赔率算出这匹马获胜的概率，并算出赌博公司总收益率；

2. 利用每匹马获胜的概率算出它夺得第一（以及第二、第三、第四）名的概率；

3. 根据50、25、10、5的计分方法估计每匹马的得分；

4. 根据对某匹马的得分的估计确定其期货价格，得分的估计值等于买入价和卖出价的平均值。买入价和卖出价的差价可以通过赌博公司的总收益率得出。

在这种情况下，所有马的得分总和为50+25+10+5=90。各匹马的买入价之和一定大于90，而各匹马的卖出价之和一定小于90。如果出现各匹马的买入价之和小于90（或者各匹马的卖出价之和大于90）的情况，可以肯定经营者犯了愚蠢的错误，你有机会利用他们的愚蠢狠狠地搂一笔钱，方法很简单：买入（或卖出）所有的马。

美国的橄榄球比赛是差额赌博的主要对象，最流行的下注方式有两种：赌两支球队得分的差额以及赌两支球队得分的总和。根据统计，一场比赛的总得分的平均值略大于40，所以期货价格通常在37～40到44～47之间。在1997年的一个周末，全部的12场橄榄球比赛的总得分数的期货价格为500～510。粗看起来这个价格的差价太小了，12场比赛的差价才达到10，而单场比赛的差价就是3。不过这个差价是经营者根据统计资

236

料得出的，他们认为这是最合理的差价。

买进价和卖出价的平均值为505，除以12得到每场比赛的总得分的期望值——42。这个值恰好是这个赛季中每场比赛的总得分的平均值。这是巧合吗？不，这是经营者精心设计的结果。期货价格中的差价由经营者的利润的变异性决定，变异性越大，则差价越大。附录3介绍了计算变异性的方法。粗略地说，几个量的和的变异性等于每个量的变异性的和，而标准差等于变异性的平方根。如果12场比赛中每场比赛的总得分数的变异性相同，那么12场比赛的总得分数的变异性就是每场比赛的总得分数的变异性的12倍，而12场比赛的总得分数的标准差就是每场比赛的总得分数的标准差的倍，约等于3.46。由于单场比赛的期货价格的差价为3，所以12场比赛的期货价格的差价为33.46，约等于10。

事实上这12场比赛的总得分数为497，因此每个卖出的客户赚了钱（收入为下注额的3倍），而每个买进的客户赔了钱（损失为下注额的13倍）。如果实际的总得分数为490，则每个卖出的客户收入下注额的10倍，而每个买进的客户损失下注额的20倍。

为了吸引客户的兴趣，经营者允许客户对在某一天（或某一个周末）发生的多个事件综合下注。1997年底的一个例子：许多场比赛同时在电视上直播，一个经营差额赌博的公司允许客户对这些比赛下注，下注的对象是以下六个量的总和：

1. 雷斯特队在橄榄球联赛中的得分数；
2. 路易斯在拳击赛中获得的点数的总和；
3. 米德布拉夫队在足球赛中的进球数的20倍；
4. 查尔萨队在另一场足球赛中的进球数的20倍；
5. 堪萨斯队在美国橄榄球赛中的得分数；

237

6. 戴维斯·拉瓦在高尔夫球赛的最后一局的杆数的 10 倍。

经营者提供的期货价格为 155～165，实际上以上六个量分别为 22、2、20、40、14 和 40，总和为 138。在这六场比赛中的前三场结束时，已得出结果的三个量的总和为 44，期货价格下调为 124～130。由于结果的不确定性下降，所以期货价格中的差价也下降。

如果你盼望经营者犯一次愚蠢的错误，并乘机捞一把，这种幻想是不切实际的。不过经营者涉足的领域非常广，并且急于扩大对客户的吸引力，所以可能给出错误的期货价格。如果你拥有丰富的体育知识，你可能抓住经营者的破绽。在 1996 年欧洲足球锦标赛中，共有 31 场比赛。经营者提供的一个下注对象是在 31 场比赛中最早的进球发生的时间，期货价格为 230～245（单位为秒）。

如果你决定卖出，你赚钱的概率极高。在全部 31 场比赛中前 230 秒内没有进球的概率相当于在一场比赛中前 118 分钟没有进球的概率，这个概率值是非常非常低的。事实上，有几场比赛的进球时间非常早，在两分钟以内。不过我是一个远离风险的人，虽然我明知赚钱的把握很大，我还是没有下注。　238

第十二章　体育生涯

如果一场比赛中的双方实力悬殊，看比赛的人会觉得索然无味。体育运动的吸引力在于结果的不可预知性。本章的目的是说明关于可能性的知识如何帮助我们更好地理解体育比赛。有时候，概率知识可以帮助运动员找到最佳策略。

足　　球

足球比赛的特点是进球数较少，即使是第一流的球队，在90分钟的比赛中的进球数也通常不超过两个。因此，同其他比赛相比，足球赛的结果具有更强的不可预知性。正是出于这个原因，重要的足球赛事总是安排很多场比赛。

在足球比赛中，进球似乎是一个随机事件。我们可以设计一个简单的数学模型描述足球赛，假定比赛双方在每一时刻进球的概率是一定的。如果这个假设成立，双方的得分符合泊松分布（请参考附录2）。统计学家在对比计算结果和实际结果时发现，二者符合得非常好。当然这个模型是有缺陷的，在实

际比赛中，第一次进球会使得其后的进球容易一些，这种现象在模型中无法反映。然而，这个模型有非常好的参考价值。表12.1是根据泊松分布得出的结论。第一列数表示四支球队的平均得分，第二列数表示这些球队在比赛中得0分的概率，第三列数表示这些球队在比赛中得1分的概率，等等。根据计算，一支平均得分为2分的球队在比赛中可能得0分，其概率为14%。用n表示一支球队的平均得分，用p（x）表示这支球队在比赛中得x分的概率，则当x=n时p（x）的值最大，而且x与n的差越大，p（x）的值就越小。 239

表12.1：假设得分符合泊松分布得出的计算结果

平均得分	0	1	2	3	4或更多
0.8	45%	36%	14%	4%	1%
1.2	30%	36%	22%	9%	3%
1.6	20%	32%	26%	14%	8%
2	14%	27%	27%	18%	14%

利用表12.1的结论可以预测比赛的结果。假设有一场比赛，主队的平均得分为1.6分，客队的平均得分为1.2分，则这两支球队的各种得分情况的概率可以在表中得到。假定两支球队的得分是相互独立的，则可以很方便地算出每种比赛结果的概率：0:0的概率为（20%）×（30%）=6%，1:1的概率为（32%）×（36%）=11.5%，等等。把各种结果为平局的情况（0:0，1:1，2:2，等等）的概率加在一起，就得到结果为平局的概率——25%。

类似地，把所有结果为主队获胜的情况的概率加在一起，

就得到主队获胜的概率——48%；把所有结果为客队获胜的情况的概率加在一起，就得到客队获胜的概率——27%。计算结果与实际的统计数据符合得很好。赌博公司在确定赔率时就是利用这种方法。如果你能发明一种更精确的计算方法，你就能更准确地预测比赛结果。你很可能因此发财。

第一次进球

在足球赛中，第一次进球是很关键的。如果一支球队率先进球，对方要花费相当长的一段时间把比分扳平。如果第一次进球发生在第88分钟，对方把比分扳平的概率几乎为0。考虑这个问题：在比赛中先进球的一方获胜的概率是多少？

240

如果两支球队实力相差悬殊，强队先进球和最终获胜的概率都极高。当然也有例外。1993年，摩纳哥队比赛刚刚开始时就攻入一球，但是英国队最终以7:1狂胜。另一个例外发生在1950年的世界杯足球赛，美国队以1:0击败英国队。美国队先入一球，并且把比分坚持到终场。但是从总体上说，当双方实力相差悬殊时，先进球的一方最终获胜的概率高达95%。我们主要关心双方实力非常接近的情况。

在这种条件下，如果全场比赛只进了一个球，显然先进球的一方获胜；如果全场比赛进了两个球，由于双方的实力非常接近，所以双方得到第二分的概率相等，即先进球的一方以2:0获胜的概率为1/2，而双方以1:1打平的概率也是1/2；如果整场比赛进了三个球，则平局是不可能的。先进球的一方失败的惟一可能是对方连进两球，其概率为1/21/2 = 1/4。所以先进球一方获胜的概率为3/4，失败的概率为1/4；如果整场

比赛进了四个球，则后三个球由哪方攻入有八种可能。用F表示先进球的一方进球，用O表示对方进球，则这八种可能表示为FFF　FFO　FOF　OFF　FOO　OFO　OOF　OOO。前四种情况使先进球的一方获胜，第五至第七种可能使双方打平，最后一种情况使先进球的一方失败。于是，先进球的一方获胜、打平或失败的概率分别为1/2，3/8，1/8。

当整场比赛的总进球数给定时，很容易计算出各种结果的概率。如果总进球数为奇数，则平局是不可能的，先进球的一方获胜的概率高于1/2；如果总进球数是不低于4的偶数，则有胜、负、和三种可能结果，先进球的一方获胜的概率为1/2。计算方法很简单，请参考附录2。

为了得到最终的结论，我们需要知道总进球数为1，2，3……的概率，总进球数为0的情况已被排除。利用泊松分布很容易算出各个概率值，当然，我们最后需要排除总进球数为0的情况。

241

在职业足球比赛中，一场比赛的得分数的平均值在2至3.5之间，2.8是一个典型的平均值。当平均得分数取2至3.5之间的值时，计算结果是一致的：排除0:0的情况，双方打成平局的概率接近于20%。所以我们可以得出一个结论：如果双方实力相当，而一方先得一分，则结果是平局的概率为20%。

平均得分数越低，第一个进球就越重要。如果平均得分数为2，则先进球一方获胜的概率为72%，由于平局的概率为20%，其失败的概率仅为8%；如果平均得分数为3.5，则先进球一方获胜的概率为64%，失败的概率为16%；如果平均得分数为2.8，则先进球一方获胜的概率为67%，失败的概率为13%。

以上结论的前提是双方实力相当。如果双方的实力有差距，我们可以认为先进球的一方继续得分的概率较高。把这个因素考虑进来，如果比赛的平均得分数为 2.8，则先得分的一方获胜的概率为 70%，失败的概率为 10%，平均的概率为 20%。1997 年 11 月，《独立报》统计了英国甲级联赛中的两个赛季以上的比赛结果。在进球数不为 0 的 825 场比赛中，68.7%的结果是先进球的一方获胜，20.8%的结果是平局，而 10.4%的结果是先进球的一方失败。从统计学的角度看，理论和实际符合得很好。这说明泊松分布可以用来预测足球比赛的结果。

在统计资料的过程中，《独立报》发现只有一支球队保持了一项可敬的纪录：这支球队在甲级联赛中出场 200 多次，多次领先，而且从未出现过这支球队先进球但最终失败的情况。你知道这是哪支球队吗？

黄　牌

黄牌是对犯规行为的警告。通常，一个队员在收到一张黄牌以后会老实很多，因为在一场比赛中累积两张黄牌就要被罚下场。如果一个队员在不同的比赛中累积了多张黄牌，他将受到停赛的处分。也许停赛对个人不算什么，但对整个球队的影响很大，因为教练的选择余地受到很大限制。有时候，一个教练发现自己的许多队员已被停赛，不得不调一个打前锋的人当后卫。如果一个队员在一场比赛中吃黄牌的概率为 p，则平均将被停赛多少次？

242

在世界杯和其他重大赛事中，通常的规定是累积两张黄牌

停赛一场，而在停赛以后黄牌数归0。这说明，一个队员身上的黄牌数有三种可能：0、1或2。当黄牌数为2时，停赛一场，而后2变成0。数学家已经准备好了一个处理这类问题的方法，即所谓的“马尔科夫链”。[①]下一章我们将讨论马尔科夫链。不过即使不用马尔科夫链的方法，这个问题也不难解决。

黄牌数从0变成1的概率与从1变成2的概率相等，所以这个队员的黄牌数为0的概率与黄牌数为1的概率相等。黄牌数从1变成2的概率为p，所以黄牌数为2的概率是黄牌数为1的概率的p倍。由于三个概率值的和为1，所以可以得出结论：黄牌数为0和1的概率都是1/（2+p），而黄牌数为2的概率是p/（2+p）。在整个赛事中，这个队员被停赛的场次所占的比例为p/（2+p）。对于一个谨慎的队员而言，p=10%，平均他有1/21的比赛被停赛。如果一个比较“猛”的队员平均每两场吃一张黄牌，则p=50%，平均他有1/5的比赛被停赛。教练应当对这些问题心中有数。

红　牌

如果防守队员的犯规行为破坏了对方的一次明显的得分机会，则这名队员将收到红牌。红牌是比黄牌更严厉的惩罚，除了下一场比赛此队员被停赛以外，在本场比赛中此队员立即出局，自己的队伍将在少一名队员的劣势下继续比赛。少一名队员当然会产生严重的后果，但是如果对方的进攻已经对球门产

① 马尔科夫链是事件的一种概率模型，其中一个事件的概率完全取决于它前面的事件。——译者注

243 生了严重的威胁，为了破坏对方的进攻，有时候吃红牌也是值得的。教练的职责之一就是教会队员如何“聪明而有效”地犯规。如果从战略上说犯规对本队有利，队员就应当毫不犹豫地犯规，而且要犯出“职业水准”。G·里德和两名同事研究了很多关于红牌的统计数据，他们的结论对足球教练有指导意义。

他们研究了荷兰的340场足球比赛，这340场比赛的共同特点是在每场比赛中有一方恰好被罚下一人。他们的目的是发现队员被罚下场的时间与比赛结果之间的关系。队员被罚下场的时间越早，则对方的成绩越好。（当然例外肯定存在，10个人打败11个人的情况也非罕见。）他们计算出对方在比赛中的平均净胜球数，并根据泊松分布算出最终对方胜出n个球的概率（n=0，1，2……），事先假定比赛双方实力相当。如果你的目的是使自己的队伍失败的概率最小化，你可以借鉴他们的结论。

一名队员是否应当犯规取决于两个因素：比赛已经进行了多少时间，以及如果你不犯规，对方进球的概率有多大。比赛已经进行的时间越短，红牌的后果就越严重；对方的进攻威胁越大，犯规就越有价值。你需要在这两个因素中作出均衡的考虑。给定对方进球的概率，你可以确定一个“标准时间”。如果当前比赛已经进行的时间少于标准时间，你不应该犯规；如果当前比赛已经进行的时间多于标准时间，你就应该犯规。下面是里德的几个典型的结论：

对方得分的概率	标准时间
100%	16分钟
60%	48分钟
30%	71分钟

以上结论针对于禁区外的犯规。如果犯规发生在禁区内，除了红牌以外，还要加罚点球。由于点球的命中率极高，所以禁区内的犯规代价很大。尽管如此，有时候犯规仍然是值得的。在临近终场时，如果你发现不犯规将导致你方必败无疑，你还有什么好犹豫的？请注意，临近终场是一个重要的条件。如果距离终场还有 1 分钟，场上比分为平局，对手马上就要进球，你可以不考虑红牌加点球的惩罚；如果距离终场还有 15 分钟，你最好老老实实地踢下去。

以上分析假定双方实力接近。如果你方的实力明显较弱，你可以更主动地犯规；如果你方的实力较强，最好稳中取胜，何必自找麻烦？

高尔夫球

多数职业锦标赛设 72 个球洞。顶尖选手之间的实力非常接近，如果一名选手在某一个球洞上比标准杆数多用了 4 杆，将对比赛结果产生很大的影响。但是在对抗赛中，由于不计总杆数，所以在某一个球洞上耗费很多杆的后果并非特别严重。例如，如果对手在一个球洞上用了 4 杆，那么你花费 5 杆或 20 杆没什么差别。在莱德杯欧美对抗赛中，每方有两名队员。在一个球洞上，哪一方花费的杆数较少，哪一方就赢得这个球洞。赢得球洞较多的一方获胜。在这种规则下，水平稳定的选手占优势。

用四个参数评价一名选手的水平。A：比标准杆数少用两杆进球的概率；B：比标准杆数少用一杆进球的概率；C：用

标准杆数进球的概率；D：用比标准杆数多的杆数进球的概率。比较两名球员 X 和 Y 的水平，两人的参数如下：

	B	C	D	
X	5%	10%	80%	5%
Y	10%	20%	40%	30%

他们在一个球洞上花费的平均杆数相同，但 X 的水平显然比 Y 稳定。在对抗赛中，平均而言 X 占优势。

如果 X 和 Y 进行一场对抗赛，我们可以算出在对每个球洞的争夺中双方获胜的概率。对于一个球洞而言，双方花费的杆数相同的概率为 36%，X 花费杆数较少的概率为 35.5%，Y 花费杆数较少的概率为 28.5%。X 占微弱优势。平均而言，在 18 个球洞中，X 将比 Y 多赢得 1 个。

假设一支球队由两名选手组成，选手的技术水平要么与 X 相同，要么与 Y 相同。则一共有三种可能：两名选手的技术水平都与 X 相同；两名选手的技术水平都与 Y 相同；一名选手的技术水平与 X 相同，而另一名选手的技术水平与 Y 相同。则整个队伍的技术水平如下：

队员构成	A	B	C	D
X—X	9.75%	18%	72%	0.25%
Y—Y	19%	32%	40%	9%
X—Y	14.5%	26%	58%	1.5%

当两个水平不稳定的选手配对时（Y—Y），整个队伍在一个球洞上花费的杆数超过标准杆数的概率从30%降低到9%。比较这三种配对方式，两个水平不稳定的选手配对要比另外两种配对方式好。当Y—Y与X—X对抗时，在一个球洞上双方打平的概率为36%，前者获胜的概率为41%，后者获胜的概率仅为23%。当Y—Y与X—Y对抗时，在一个球洞上前者获胜的概率为36%，后者获胜的概率为30%。当X—Y与X—X对抗时，在一个球洞上前者获胜的概率为32%，后者获胜的概率仅为20%。

另外一种比赛方式也对水平稳定的选手有利。假设多名选手一同参赛，每个人拿出一笔钱凑成一笔奖金，奖金被分配到各个球洞上。几个人轮流击球。在一个球洞上，如果有一个人用的杆数比其他人都少，则此人赢得这个球洞上的奖金；如果没有决出一个最强者，则奖金累积到下一个球洞。假设5个人参赛，其中四个人水平稳定，而一个人水平不稳定，则这个家伙处于非常可怜的地位。通常他只能期待在自己把球击飞情况下另外几个人分不出高下。

在其他体育比赛中，水平的稳定性也是一个重要的因素。在跳远和投掷项目中，一个选手可以尝试多次，但只取最好成绩。有两个链球选手p和q，p的平均成绩为75米，变异性为16；q的平均成绩为77米，变异性为4。q的平均成绩比p好，但在比赛中p获胜的机会更大，因为q的成绩比较稳定。

246

如果链球的比赛规则改为一掷定胜负，则比赛对q有利，其获胜的概率高达67%。通常比赛规定每人掷六次，取最好成绩。此时情况对p有利，其获胜的概率为58%。如果规定每人掷20次，则p获胜的概率高达75%。显然，p打破世界纪录的概率要比q高很多。

板　　球

多年以来，统计学在板球中的应用一直停留在统计得分、计算平均值的水平上，偶尔为教练提供一点参考信息。但是最近，板球专家开始注重有效地运用统计学技术。弗兰克·德克华兹和汤尼·路易斯创造性地解决了一个30年来悬而未决的重要问题：如果一场比赛因下雨等意外因素暂停，如何根据当前的比分判断胜负？

考虑一个典型的例子：比赛规定每方击50次球，A队已经击完50次球，得到200分，而后因下雨比赛暂停。雨停之后，剩下的时间只够B队击25次球。应当要求B队在25次击球中得多少分？

答案可以有很多种。最简单的答案是要求B队在25次击球中的得分达到A队的得分的一半。B队如果超过这个目标则获胜，如果未完成这个目标则失败。当然这个答案很不准确，因为在后半场中得分更加困难，A队吃亏了。另外一个答案是取出A队最成功的25次击球，统计出这25次击球的得分，要求B队达到这个分数。但是这样对B队又是不公平的。有一个真实的例子：南非队在一场比赛中占据优势，他们只要在13次击球中得到22分就可以获胜，对于一支职业队来说，这是轻而易举的任务。但此时下雨使比赛暂停，比赛恢复以后剩下的时间只够击球1次，而南非队必须在1次击球中得到21分才能获胜！

德克华兹和路易斯的主要思路在于，除了剩余的击球次数以外，球场上还剩下多少个球门也是需要考虑的重要因素。他

们分析了许多年的比赛积累下来的数据，得出一个数学模型，根据剩余的击球次数和球场上剩余的球门数确定一支球队的胜利条件。考虑这个例子：印度队击球 50 次得 250 分，巴基斯坦队击球 20 次得 100 分，双方得分的速度相同。此时比赛因下雨暂停，雨停之后，剩下的时间还够巴基斯坦队击球 10 次，如何确定巴基斯坦队的胜利条件？

为了使这个获胜条件对双方公平，必须考虑在比赛暂停时球场上还剩下多少个球门。如果全部 10 个球门都在，则巴基斯坦队的早期表现非常好。如果没有下雨干扰比赛的话，他们在剩下的 30 次进球中得 251 分以上的机会相当大。如果比赛暂停时他们已经损失了 3 个球门，则双方获胜的机会相当；如果比赛暂停时他们损失的球门多于 3 个，则他们已经居于劣势。在德克华兹和路易斯的模型中，这三种情况需要分别处理。而在以往的处理方法中，不考虑球场上剩下的球门数，仅规定巴基斯坦队需要在 10 次击球中得到 50 分。如果消耗的球门数已达到 3，这个规定对巴基斯坦队有利；如果消耗的球门数已达到 6，这个规定足以使巴基斯坦队在绝望中起死回生。

德克华兹和路易斯的主要技巧在于计算一个参数："得分机会"，这个参数是由已经消耗的击球数和球门数决定的。如果在 20 次击球之后，场上全部 10 个球门都在，则这个球队的得分机会为 77.1%；如果已经消耗了 3 个球门，则他们的得分机会只剩下 62.3%；如果已经消耗了 3 个球门，则他们的得分机会只剩下 62.3%；如果已经消耗了 6 个球门，则他们的得分机会只剩下 36.2%。由于下雨的影响，现在他们只剩下 10 次击球机会。在第一种情况下（消耗的球门数为 0），巴基斯坦队当前的得分机会为 34.1%，下雨使他们损失的得分机会为 77.1% - 34.1% = 43%。250 分（对方的得分）的 43% 为

248

107.5，刨出这个分数，他们需要取得 250 - 107.5 = 142.5 分。由于他们已经得到了 100 分，所以只要在剩下的 10 次击球中得到 42.5 分就可以了。最终他们如果得到 43 分则获胜，如果得到 42 分则失败，平局是不可能的。这个目标对于他们来说是不难完成的。他们之所以拥有比较舒服的获胜条件，是因为他们在前 20 次击球中表现出色：他们完成了这个目标（得 250 分）的 40%，但只消耗了 22.9% 的得分机会（100% - 77.1%）。

在第二种情况下（消耗的球门数为 3），巴基斯坦队当前的得分机会为 31.4%，下雨使他们损失了 62.3% - 31.4% = 30.9%的得分机会。250 分的 30.9%为 77.25，所以他们需要得到 172.75 分。由于他们已经得到 100 分，所以在 10 次击球中再得 73 分就可以获胜。在比赛暂停时，他们完成了全部目标的 40%，而消耗了 100% - 62.3% = 37.7%的得分机会，所以他们的表现与对手相仿。但如果他们消耗的球门数不止 3 个，则他们在暂停以前的表现不如对手。

在最后一种情况下（消耗的球门数为 6），巴基斯坦队当前的得分机会为 24.6%，下雨使他们损失了 36.2% - 24.6% = 11.6%的得分机会。250 分的 11.6%为 29，所以他们需要得到 221 分。由于他们已经得到 100 分，所以在 10 次击球中需要再得 121 分。恰好得 121 分则为平局，超过 121 分则获胜，不足 121 分则失败。当然，在 10 次击球中得 121 分是很困难的。这是因为他们在比赛暂停以前消耗了太多的得分机会。

所有这些结论基于 1997 年以前的统计数据。随着数据的更新，结论可能需要进行一些调整，但是基本思路是不变的。

德克华兹和路易斯的模型可以应用于比赛的任何阶段。假设英国队对抗澳大利亚队，比赛原计划双方各击球 50 次，英

国队先击球。当英国队击球30次之后，开始下雨，比赛暂停，此时英国队的得分为154。恢复比赛时，剩下的时间只够击球30次。显然，不应该让英国队继续击球。英国队的击球已经249 结束，下面轮到澳大利亚队击球。英国队被剥夺了20次击球机会，当时剩下8个球门。英国队可能认为他们的开局打得很好（还剩下很多球门），如果规定对方的获胜条件是在30次击球中得分154以上，对他们是不公平的。事实上德克华兹和路易斯的模型也这么认为。当比赛暂停时英国队还有54%的得分机会，这些得分机会已被剥夺。澳大利亚队被剥夺了前20次击球。当他们开始击球时，他们的得分机会只剩下77.1%，因此，澳大利亚队被剥夺了22.9%的得分机会。相比之下，英国队的损失更大，其差额为54% - 22.9% = 31.15%。根据多年比赛的统计结果，在50次击球的比赛中的平均得分为225，所以英国队的得分应当增加225的31.15%，即69.975分，英国队的得分应被调整为154 + 69.975 = 223.975。结论：如果澳大利亚队在30次击球中得到224分或更多则获胜，否则失败。[①]我认为板球队员很难接受如此复杂的计算。不过政治家的箴言是“不受欢迎的决定通常是正确的决定”，类似地我们可以说：“既然德克华兹和路易斯的模型如此难懂，这个模型多半是正确的。”

最后一个例子：肯特队对抗埃塞克斯队，肯特队在50次击球中得200分。在轮到埃塞克斯队击球时开始下雨，而雨停

---

① 作者的结论似乎有问题。为什么要参照各队在比赛中的平均得分，而不参照英国队自己在本场比赛中的表现呢？在本场比赛中，英国队用46%的得分机会得到154分，所以31.15%的得分机会对应的得分应当是（154/46%）×31.15% = 104.285。对方的获胜条件应当是得分154 + 104.285 = 258.285以上。莫非他觉得258.285的结论有点高得离谱？——译者注

后预定的比赛结束时间已经过了。总不能连一次击球的机会也不给，所以允许埃塞克斯队击球 10 次。应当如何确定他们的获胜条件？裁判可以确立一个标准，如果埃塞克斯队达到了这个标准则获胜，否则失败。但是这个标准同时与他们的得分数和最终剩下的球门数有关。所以他们必须同时考虑两个因素，尽量使比分更高而消耗的球门数更少。他们的教练可能不得不拿着计算器不停地推算各种结果。

网　球

网球讲解员经常用到“比赛已经达到最关键的时刻”之类的说法。潘松·冈萨雷斯是他那个时代最优秀的网球手，他认为比分达到 15～30 时是最关键的时刻。他的解释是，当比分达到 15～30 时，如果发球方再胜一球，比分达到 30～30，发球方获胜的机会很高；如果对方再胜一球，比分达到 15～40，发球方获胜的机会就很低了。其他的网球手和教练提出了不同的比分作为最关键的时刻。统计学家卡尔·莫里斯最终解决了这个问题，并给出了严格的证明。

莫里斯认为，评价一个局面的重要程度的标准是两个概率值的差。首先，计算出在此局面下如果发球方获胜，他最终获胜的概率值 x；其次，计算出在此局面下如果发球方失败，他最终获胜的概率值 y。x 和 y 的差就可以反映这个局面的重要程度。如果某个局面使得这个差达到最大值，则这个局面就是最关键的局面。显然，如果一个局面对于发球方而言是最关键的局面，则这个局面对于接球方而言也是最关键的局面，反之亦然。根据莫里斯的标准，冈萨雷斯的结论肯定是错误的，因

为有一个局面总是比场上比分为 15～30 的局面更重要，即场上比分为 30～40 的局面。

比较一下这两个局面。当场上比分为 15～30 时，如果发球方再赢一球，则比分变成 30～30，此时发球方的获胜条件是超出对手两球，用 a 表示这个概率值。显然，这个值与比分为 40～40 时发球方获胜的概率值相同（因为在此情况下发球方的获胜条件也是超出对手两球）；如果发球方再丢一球，则比分变成 15～40，此时发球方获胜的概率已经很低了，但概率值毕竟不为 0。所以这两个概率值的差应当小于 a。

再看场上比分为 30～40 的局面。如果发球方再赢一球，则比分变成 40～40，此时发球方的获胜条件是超出对手两球，其概率值也是 a；如果发球方再丢一球，则对手已经获胜，发球方获胜的概率为 0。所以这两个概率值的差为 a。比较两个差额，结论是第二种局面更重要。

莫里斯发现，没有一种简单的通用原则用来判断哪种局面更重要。在某些情况下，一个局面的重要性取决于发球方赢得一球的概率值 p。当 p 大于 50%时，可以确定最关键的局面是比分为 30～40 的时候。当 p 在 50%到 62%之间时，次重要的局面是双方以高比分打平的场合（即比分为 30～30，40～40 等等的局面）；但当 p 大于 62%时，次重要的局面是比分为 15～40 的局面。最不重要的局面是比分为 40～0 的场合。计算的过程不复杂，但是需要考虑很多因素。 251

另外一个问题。发球方在发球时可以选择两个方案：发快球或发慢球。发快球的优点在于得分机会较高，发慢球的优点在于发球违例的概率较低。通常的策略是先发一个快球，如果第一次发球违例，则再发一个慢球。这种策略是发球方的最佳策略吗？以得分概率最大化为目标，最佳策略是什么？

S·L·乔治研究了这个问题。他认为，有四个因素与最终的结论有关：

○ 发快球不违例的概率，计为 x；

○ 发慢球不违例的概率，计为 y；

○ 在发快球不违例的前提下得分的概率，计为 F；

○ 在发慢球不违例的前提下得分的概率，计为 S。

假定 $x>y$，$F>S$。发球方可以选择的发球策略有四种：先发快球，如果发球违例再发慢球；总是发快球；总是发慢球；先发慢球，如果发球违例再发快球。我们可以计算出在各种策略下发球方得分的概率。

在第一种策略下，发球方有两种得分机会：

1. 发快球不违例，并且得分，其概率为 xF；

2. 发快球违例，随后发慢球，不违例并得分，其概率为 $(1-x)yS$。

因此，在第一种策略下得分的概率为 $xF+(1-x)yS$。类似地可以算出另外三种策略对应的得分概率。

一个明显的结论是第一种策略总是比第四种策略好，因此，第四种策略已经可以排除，只需要考虑前三种策略。在这三种策略的比较中，优劣取决于具体的 x、y、S 和 F 的值。最重要的标准在于 xF 与 yS 的比值。xF 是通过发快球直接得分的概率，yS 是通过发慢球直接得分的概率。用 R 表示 xF/yS。

252

通常，R 小于 1。但是如果 R 大于 1，则每次都应当发快球。这是因为 R 大于 1 意味着在任何情况下发快球的结果都优于发慢球，即使第一次发球已经违例。在职业高手之间，R 大于 1 的情况很罕见，所以职业高手的常见选择是如果第一次发球违例则发慢球。

当 R 小于 1 时，第一种策略最佳。但也有例外。当 R 值

充分小时（即发慢球直接得分的概率很高时）最佳策略是每次都发慢球。什么情况下我们可以认为 R 值充分小呢？条件是 R 小于 $1+x-y$。当然，在职业高手之间这种情况很少见。当 $x=50\%$，$y=90\%$，$F=80\%$，$S=50\%$ 时，$xF=40\%$，$yS=45\%$，$R=xF/yS=40/45<1$，此时应当选择第一种策略。但是，如果 s 只有 40%，则 $yS=36\%$，$R=40/36>1$，此时最佳策略是每次都发快球。

由于某些局面比其他局面更重要，所以一名选手可以考虑在重要局面下更加卖力，而在不重要的局面下节省一点体力。在第八章的研究中，我们假定一名选手赢得一球的概率值是固定的，现在我们假定选手可以选择是否在对某一球的争夺中投入较多的体力。是否可以保证赢得本局比赛的概率不变，但是消耗的体力较少？

我们的模型不够完美，但有很好的参考价值。考虑一名选手斯密斯在比赛中的行为。在对每一球的争夺中，斯密斯可以选择的策略有两种：耗费较多的体力以增加得分的概率；节省体力放弃一部分得分的概率。一般性的原则是在比较重要的局面下选择前者，而在其他局面下选择后者。具体地说，斯密斯在三种情况下应当选择节省体力的策略：对第一个球的争夺；自己的得分领先；比分达到 0～40。在其他情况下，斯密斯应当不惜体力，争取更高的得分概率。这种战略的结果是斯密斯在比赛中选择两种策略的频率大致相等。　253

表 12.2：为了使最终获胜的概率不变，
两种策略下的得分概率

低体力损耗下得分的概率	高体力损耗下得分的概率
58.5%	61%
57%	62%
55%	63%
53%	64%
50.5%	65%

比较两个模型。在第一个模型下，斯密斯始终以相同的体力消耗争夺每一个球，每一次得分的概率都是 60%。在这种情况下，他最终获胜的概率为 73.5%。在第二个模型下，在对每一个球的争夺中，他可以选择自己投入的体力，投入的体力较多使得得分的概率上升。可选择的策略有两种，选择方式如上所述。为了使斯密斯最终获胜的概率达到 73.5%，在两种体力损耗标准下的得分的概率分别是多少?

结论如表 12.2 所示。如果斯密斯按照我们规定的战略应用两种策略，则他最终获胜的概率也是 73.5%。表中的最后一行数据表明，即使斯密斯在节省体力的策略下的得分概率只有 50.5%，他仍然可以保证最终获胜的概率为 73.5%。利用这种策略不改变获胜的概率，但可以在总体上节省体力。

锦标赛的比赛规则

锦标赛的目标是在参赛选手（或队伍）中决出最强者。当

然，最终的冠军未必是实力最强者。如果你相信实力最强者一定会夺冠，本节就没有讨论的必要。实际情况是，不同的比赛规则会影响实力最强者夺冠的概率。不过，有时候比赛的目的不仅是决出冠军，我们还希望欣赏一定数量的比赛。在世界杯足球赛中，参赛队有 32 支，如果在比赛的第一天就把一些球队送回家显然太匆忙了。

254

锦标赛中常见的规则有很多种。在简单循环制中，每两个参赛者恰好相遇一次。在足球联赛中，每两支参赛队伍相遇两次，每支球队充当一次主队。在淘汰制中，通过抽签随机决定配对方案。网球赛通常采用淘汰制，但安排了种子选手，在比赛早期保护种子选手。在象棋赛和桥牌赛中，如果参赛选手较多，通常采用积分循环制：首轮比赛抽签决定，而后安排积分接近的选手对抗。槌球赛有时使用抽签循环制，在比赛早期避免出现三支球队之间的循环比赛。如果最后出现两支积分相同的球队，则二者进行决赛。

戴维·阿帕莱顿用统计学方法比较了几种赛制。他为每个参赛者分配了胜率，实力较强的队伍胜率较高。当两个参赛者相遇时，根据双方的胜率判断每一方取胜的概率。他摹拟了当参赛者为 8 名或 16 名时各种赛制的结果，每种赛制摹拟数千次，以此判定各种赛制的优劣。具体的结论与参赛人数、每个参赛者的胜率等因素有关，但戴维得出了一些普遍性的结论。

我们的目标是发现一种使胜率最高的选手最终获胜的概率最大的比赛规则。在戴维比较的八种赛制中，最好的赛制是每两名选手相遇两次的循环赛。这种赛制要求的比赛场次的数目也是最多的。排在第二位的赛制是安排种子选手的抽签循环制。这种赛制的优点有二：其一，胜率最高的选手作为种子选手受到保护，最终胜出的概率较高；其二，胜率最高的选手有

两种进入最后决赛的机会。排在第三位的是三局两胜的淘汰赛。在这种赛制下，如果胜率最高的选手偶然失手一次，他还有第二次机会免遭淘汰。排在第四位的是不安排种子选手的抽签循环赛。排在后面的四种赛制依次为简单循环制、安排种子选手的淘汰制、进行五或六轮的积分循环制和不安排种子选手的淘汰制。积分循环制和循环赛的缺点是可能出现两名选手并列第一的结果，其他赛制没有这个问题。

255

表 12.3

A	B	C	D	
A		W	W	L
B	L		W	W
C	L	L		W
D	W	L	L	

如果出现积分相同的情况，有若干种排序方法。假设有四名选手参加循环赛，每场比赛都分出胜负，比赛结果如表 12.3 所示，表中 W 表示获胜，L 表示失败。A 和 B 获胜两次，C 和 D 获胜一次。如何决定四个人的次序？一种办法是参照同分者之间的比赛结果。由于 A 胜 B 而 C 胜 D，所以最终次序为 A、B、C、D。不过 D 会觉得委屈，他战胜了冠军，却名列最后。

象棋比赛中经常使用另一种方法：比较对手分。由于 A 战胜了 B 和 C，所以 A 的对手分是 B 和 C 的得分的总和，即 3 分。类似地，B、C、D 的对手分分别是 2 分、1 分和 2 分。A 的对手分最高，但是 B 和 D 的对手分相同。不要紧，我们以

对手分为依据再计算第二级的对手分，四个人的分数分别是3、3、2、3。现在还是分不出次序，不过再重复两次就可以分出次序，四人的得分分别是8、6、5、3。如果继续进行这个程序，四个人的得分值会有变化，但次序始终保持不变：A、B、D、C。这种方法看起来很公平。[①]

在正规比赛中通常事先规定好了处理积分相同的情况的方法。不过在友谊赛中可能没有事先规定。此时你可以随便选一种办法。

256

① 作者似乎把问题想复杂了。象棋比赛的实际规则是仅比较同分者的对手分。A与B同分，但A的对手分较高，所以A优于B；C与D同分，但D的对手分较高，所以D优于C。B与D的对手分相同，但是B的积分高于D，所以根本不需要比较他们两个的对手分。因此，仅计算一次对手分就可以决定次序。——译者注。

第十三章　其他需要运气的问题

正如浮在水面以上的仅仅是冰山的一角，在这本书中我们只能涉及概率知识的一小部分。在最后一章里，我们讨论一些简单的概率问题。

扑克牌问题

在惠斯特[①]和桥牌中，一手牌由 13 张牌构成，所以每一手牌有

$$^{52}C_{13} = 635\ 013\ 559\ 600 \text{ 种可能}$$

根据我们在第七章得出的结论，我们每玩 100 万局桥牌，遇到两手完全相同的牌的概率为 50%。我们可以认为这个频率等于 0。如果考虑全部四个人手中的牌，将得到一个天文数

---

① 一种扑克游戏，桥牌的前身——译者注

字。无论你手中得到什么牌，你的上家手中的牌有$^{39}C_{13}$种可能，而你的下家手中的牌有$^{26}C_{13}$种可能。在三个人手中的牌决定以后，第四个人手中的牌就确定了。四个人手中的牌有多少种可能？答案是一个 29 位数：

53 644 737 765 488 792 839 237 440 000

最精确的叫牌体系也无法覆盖每一种情况。

根据概率知识你可以知道拿到某一类牌的概率。例如，你有一个缺门的概率是多少？某一个花色有 8 张的概率是多少？手中的大牌有 25 点的概率是多少？你的牌型为 4—3—3—3 或 4—4—3—2 的概率是多少？当对方手中有一张 Q 时，你应当硬敲还是飞牌？等等。[1]

首先考虑一个简单的问题：你的牌型为 4—4—3—2 的概率是多少？第一步，算出你持有 4 张黑桃、4 张红桃、3 张方块、2 张草花有多少种可能。利用附录 1 中的知识，可以得出结论：$X = {}^{13}C_4 \times {}^{13}C_4 \times {}^{13}C_3 \times {}^{13}C_2$。此外，需要考虑 4 种花色的排列。在 4 种花色中选 2 种作为 4 张套，有$^4C_2 = 6$种可能，在剩下的 2 种花色中选 1 种作为 3 张套，有 2 种可能，最后剩下的 1 种花色只能作为 2 张套。所以牌型为 4—4—3—2 的可能种类一共有 $6 \times 2 \times X = 12X$ 种。以前我们已经算出一手牌一共可能有多少种，这两个值的比就是牌型为 4—4—3—2 的概率。

257

① 这里涉及几个桥牌术语。按照通行的桥牌理论，A、K、Q、J 被当做“大牌”，其“点数”分别为 4、3、2、1。如果你持有四张 A 和三张 K，没有 Q 和 J，你手中的大牌点就是 25。“4—3—3—3”指手中的四门花色分别为 4 张、3 张、3 张、3 张。“硬敲”指直接从手中出大牌，希望对方手中的大牌自动掉出来。硬敲的前提是对方的大牌没有其他牌保护。“飞牌”指从没有大牌的一方出小牌，如果下一家跟小牌就放过，如果下一家出大牌就用大牌盖住。飞牌的前提是对方的大牌位于你方的大牌的上方。——译者注

利用类似的方法可以计算出牌型为 w—x—y—z 的概率。主要计算过程为两步：首先计算持有 w 张黑桃、x 张红桃、y 张方块、z 张草花有多少种可能。其次，为第一步的结论寻找一个适当的系数。如果 w、x、y、z 这 4 个数各不相同，则系数为 24；如果在 w、x、y、z 这 4 个数中有 2 个相同，则系数为 12；如果在 w、x、y、z 这 4 个数中有 3 个相同，则系数为 4。

总共有 39 种可能的牌型，出现的频率最高的牌型是 4—4—3—2。在这 39 种牌型中，只有 13 种牌型出现的频率高于 1%，如表 13.1 所示。表 13.1 的最后一列是第二列的累积结果。表中未列出的 26 种牌型出现的频率总和不足 1/20。许多玩家认为，牌型越平均，则出现的频率越高，所以 4—3—3—3 牌型出现的频率应当最高，但事实上 4—3—3—3 的频率仅排在第五位，其频率为 10.54%。表 13.1 的结论基于一个假设：洗牌过程是严格的。在实战中，洗牌过程往往比较匆忙，所以实战结果可能与表 13.1 的结论有出入。在正式比赛中，往往用计算机替代人工洗牌，在这种情况下，表 13.1 可以提供更好的预测。

除牌型以外，我们还可以算出长套、缺门和单张出现的概率。（长套指手中张数较多的花色，缺门指手中没有某一花色的牌，单张指某一花色的牌手中只有一张。）有三种牌型使得手中最长套仅有 4 张（即 4—4—3—2，4—3—3—3 和 4—4—4—1），所以这三种牌型的概率之和即最长套为 4 张的概率。类似地可以算出最长套的张数为 5、6、7、8 或更多的概率。结论在表 13.2 中。持有缺门的概率为 5%，持有单张但没有缺门的概率为 31%，持有两个缺门的概率为 1/10000。

表 13.1

牌型	出现的概率	累积值
4—4—3—2	0.215 5	0.215 5
5—3—3—2	0.155 2	0.370 7
5—4—3—1	0.129 3	0.500 0
5—4—2—2	0.105 8	0.605 8
4—3—3—3	0.105 4	0.711 1
6—3—2—2	0.056 4	0.767 6
6—4—2—1	0.047 0	0.814 6
6—3—3—1	0.034 5	0.849 1
5—5—2—1	0.031 7	0.880 8
4—4—4—1	0.029 9	0.910 7
7—3—2—1	0.018 8	0.929 6
6—4—3—0	0.013 3	0.942 8
5—4—4—0	0.012 4	0.955 2

表 13.2

最长套的张数	出现的概率	累积值
4	0.350 8	0.350 8
5	0.443 4	0.794 2
6	0.165 5	0.959 7
7	0.035 3	0.994 9
8 或更多	0.005 1	1

必然存在一个花色，你和你的同伴至少持有其中 7 张。存在一门花色你和你的同伴至少持有其中 8 张的概率接近 85%。

存在一门花色你和你的同伴至少持有其中 10 张的概率接近 10%。计算这类问题的方法与表 13.1 类似，但是我们考虑的不是 13 张牌，而是 26 张牌。把你的牌与同伴的牌合在一起计算，出现的频率最高的牌型为 8—7—6—5，频率接近 1/5。对方的情况也是如此。

4 个人每人持有 1 张 A 的概率在 1/9 到 1/10 之间。4 张 A 在 4 个人手中的分布以 2—1—1—0 最为常见，频率高达 58.4%。出现的频率较低的分布情况依次为 3—1—0—0，2—2—0—0，1—1—1—1，4—0—0—0。最后一种分布出现的频率为 1%。

259

在桥牌中，最常用的衡量一手牌的价值的方法是密尔顿·沃克发明的大牌点法。A、K、Q、J 被当做大牌，其点数分别为 4，3，2，1。一副牌的点数之和为 40，平均每人持有 10 点。一手牌最少可能有 0 点，最多可能有 37 点。（除大牌点以外，长套、缺门和单张也有点数，不过我们在这本书中不予讨论。）统计各种牌点数对应的概率是一件很繁琐的事，不过我们可以让计算机完成这个任务。表 13.3 是计算机提供的答案。

表中的数据基本对称，对称中心为 9.5 点。10 点是中位数，出现的频率也最高。标准差略大于 4。一手牌的大牌点不超过 3 的概率为 5%，不低于 20 的概率小于 1%，不低于 25 点的概率仅为 1/2000。

叫牌之后开始打牌，防守方出第一张牌，这张牌称为“首攻”。首攻之后明手亮牌。此时除明手外每个人都能看到 26 张牌（自己的 13 张牌和明手的 13 张牌），现在大家开始计算看不见的 26 张牌的分布情况。叫牌过程可以提供计算的线索，单纯的算数运算也很有用。当你主打无将时，长套经常为你方提供大量的赢墩，有时候长套中的“6”和“7”和 A 一样重

要。对于庄家来说，最重要的是分析关键牌张在两个对手中的分布情况。

表 13.3

大牌点数	对应的概率	累积值
5 或更少	0.140 0	0.140 0
6	0.065 6	0.205 6
7	0.080 3	0.285 8
8	0.089 0	0.374 8
9	0.093 5	0.468 3
10	0.094 0	0.562 4
11	0.089 4	0.651 8
12	0.080 3	0.732 1
13	0.069 1	0.801 2
14	0.057 0	0.858 2
15	0.044 3	0.902 4
16 或更多	0.097 6	1

260

例如，你方持有 8 张草花，对方持有 5 张。这 5 张草花在两个对手中的分布可能是 5—0、4—1 或 3—2。通常 3—2 分布对你方最有利。实际分布为 3—2 的概率是多少？按 3—2 分布设计打牌方案与按飞牌设计打牌方案孰优孰劣？（飞牌的成功率为 50%。）为了回答这个问题，先要算出你左手方的对手持有 3 张草花的概率。对方一共有 26 张牌，其中 5 张是草花，21 张是非草花。你左手方的对手持有 3 张草花即意味着在 5 张

草花中抽出 3 张、在 21 张非草花中抽出 10 张构成他手中的牌。在 5 张草花中抽出 3 张有${}^{5}C_{3}$种可能，在 21 张非草花中抽出 10 张有${}^{21}C_{10}$种可能，所以他持有 3 张草花的可能有

$$ {}^{5}C_{3}\times {}^{21}C_{10}\text{种} $$

由于他手中的牌一共有${}^{26}C_{13}$种可能，所以他持有 3 张草花的概率为${}^{5}C_{3}\times {}^{21}C_{10}/{}^{26}C_{13}=0.339\,1$。此外，你右手方的对手持有 3 张草花的概率也是 0.339 1，所以对方的 5 张草花 3—2 分配的概率为 68%。表 13.4 提供了类似的问题的答案。

表 13.4 对于实战有很好的指导价值。根据表中的结论，3—2 分配的概率高于飞牌成功的概率，但是 3—3 分配的概率低于飞牌成功的概率。当你在叫牌过程中面临抉择时，这个表也能提供帮助。例如，你相信完成 2 无将是没问题的，你是否应当加叫到 3 无将?

表 13.4

你方手中的张数	对方的牌张分配方式	对应的概率
7	3—3	0.355
	4—2	0.485
	5—1	0.145
	6—0	0.105
8	3—2	0.678
	4—1	0.283
	5—0	0.039
9	2—2	0.407

你方手中的张数	对方的牌张分配方式	对应的概率
	3—1	0.497
	4—0	0.096
10	2—1	0.78
	3—0	0.22
11	1—1	0.52
	2—0	0.48

你的决定取决于你有多大把握完成 3 无将。在盘式桥牌中，不成局定约的价值是不确定的，这使得问题比较复杂。为了简化问题，采用复式桥牌中的计分标准。在复式桥牌的规则下，完成不成局定约的奖分为 50 分，完成成局定约的奖分为 300 分（作为无局方）。你可能遇到 4 种情况：1. 叫到 2 无将并完成；2. 叫到 2 无将但完成 3 无将；3. 叫到 3 无将但仅完成 2 无将；3. 叫到 3 无将并完成。各种情况的得分如下：

261

	完成 2 无将	完成 3 无将
叫到 2 无将	120	150
叫到 3 无将	-50	400

假设完成 3 无将的概率为 p，完成 2 无将的概率为 1 - p，则叫到 2 无将的平均得分为 150p + 120（1 - p） = 120 + 30p，而叫到 3 无将的平均得分为 400p - 50（1 - p） = 450p - 50。当 p 大于 17/42（接近于 40%）时后者较好。作为无局方，如果你完成 3 无将的希望在于对方的 3—3 分配，则你的成功率低于

40%，你叫到 3 无将是不明智的。作为有局方，情况有所不同，因为计分方法不同。如果未完成定约，作为有局方的罚分比作为无局方高一倍，但另一方面，如果完成成局定约，奖分也由 300 分增加到 500 分。利用同样方法可以算出有局方叫到 3 无将的条件：如果成功率超过 22/67（接近于 33%），叫到 3 无将就是值得的。如果你的成功同样依赖于对方的 3—3 分布，作为有局方叫到 3 无将是合理的决定。如果成功的希望在于一次飞牌（飞牌的成功率为 50%），无论作为有局方还是无局方都应当叫到 3 无将。

以上分析以平均得分为判断依据。在复式桥牌中，情况更复杂。因为你必须考虑别人在面对同样的问题时将如何抉择，你对别人的估计将决定你的最终选择。

一个类似的问题：你有把握完成小满贯，你是否应当加叫到大满贯？作为有局方时，成局奖分为 500 分，小满贯奖分为 750 分，大满贯奖分为 1 500 分。假设定约为无将，你可能遇到 4 种情况：1. 叫到 6 无将并完成；2. 叫到 6 无将但完成 7 无将；3. 叫到 7 无将但仅完成 6 无将；4. 叫到 7 无将并完成。各种情况的得分如下：

	完成 6 无将	完成 7 无将
叫到 6 无将	1 440	1 470
叫到 7 无将	- 100	2 220

计算方法完全相同。假设完成 7 无将的概率为 p，完成 6 无将的概率为 1 - p，则叫到 6 无将的平均得分为 1 440 + 30p，而叫到 7 无将的平均得分为 2 320p - 100。当 p 大于 154/229（接近

于67%）时后者较好。作为无局方结论也是如此。如果你发现完成大满贯的条件是对方的3—2分配（概率为68%），叫到大满贯是正确的。但如果成功的希望在于一次飞牌，你就不应该冒险。在6阶水平上如果你想再进一步，必须保证2:1以上的成功率。

如果在你方的长套上对方持有Q，应当如何处理？用A和K硬敲还是先打一张大牌然后飞牌？正确的策略取决于你方在这个花色中持有多少张牌。如果你方在这个花色上持有9张牌，分配为5—4，在对方的Q是单张或对方2—2分配的情况下硬敲可以获得成功，概率为53.1%。此外，如果对方4—0分配，你还有机会中途改变打法，这使得成功的机会略微增加。总的来说，第一种打法的成功率为58%。飞牌的成功率为50%，但是先打一张大牌使得成功的机会增加一些，因为你可能及时发现对方的4—0分配，也有可能抓住对方的单张Q。第二种打法的成功率为56%。这两种打法的差别不是很大，但已足以使高手作出抉择。如果你方在这个花色上持有8张牌，分配为4—4，两种打法的成功率分别为33%和51%，显然第二种打法较好。如果你方在这个花色上持有7张牌，则你有更充分的理由选择第二种打法。

262

有时你可以在打完几轮牌之后再作决定。在这种情况下，伴随着打牌过程的进行，牌张分配的概率会有变化。例如，在打完7轮牌之后，你方持有9张黑桃，Q在对方手中，对方没出过黑桃。此时，对方手中有12张牌，其中4张是黑桃，8张是非黑桃。4张黑桃在对方手中的分布为2—2、3—1、4—0的概率分别为45.5%、48.5%、6%。此时，第一种打法的成功率为60%，第二种打法的成功率为56%。在打完几轮牌以后再作决定通常比在一开始就作决定有利，不过除非进行到终局

再作决定，否则成功率不会有太大变化。

洗牌问题

常见的洗牌方法有两种：弹式洗牌和插式洗牌。弹式洗牌是把一幅牌均匀地分为两叠，然后把两叠牌混合在一起。完美的弹式洗牌要求两叠牌数目完全相等，而且每隔 1 张插入 1 张。完美的弹式洗牌又分为两种：一种方法使原来位于第 1 张的牌在洗牌之后仍然位于第 1 张，另一种方法使原来位于第 1 张的牌在洗牌之后位于第 2 张。用第一种方法洗牌的结果是第 1 张牌和最后 1 张牌的位置不变，第 18 张牌和第 35 张牌位置互换，此外的 48 张牌在 6 个区域内循环，每个区域包括 8 张牌。也就是说，一副牌经过 8 次洗牌之后又回到原来的状态。263 在实战中最好不用这种方式洗牌，因为洗牌者有机会作弊。如果用第二种方法洗牌，每洗 52 次可以使整副牌还原。

插式洗牌的过程通常是把牌一张一张地从一只手传到另一只手：先把第 1 张牌传过去，然后把第 2 张牌放在上面，把第 2 张牌放在下面，把第 3 张牌放在上面，把第 4 张牌放在下面，如此等等。完美的插式洗牌的结果是第 18 张牌位置不变，第 11 张牌和第 32 张牌位置互换，其余的牌进入 6 个区域循环，6 个区域分别包括 3、4、6、12、12、12 张牌。经过 12 次完美的插式洗牌之后，整副牌还原。

在实际玩牌的时候，很少有人耐心地使洗牌符合完美的弹式洗牌的标准，更不要说洗 8 次牌了。根据理查德·爱普生的研究，初学者洗牌通常用 0.7 秒，专家洗牌通常用 0.5 秒。初学者洗牌的时候，经常把一沓牌留在顶端或底端，而中间的牌

有时每隔 1 张插入 1 张，有时每隔 2 张插入 1 张。专家洗牌的时候，有时每隔 1 张插入 1 张，有时每隔 2 张插入 1 张，前者的频率通常是后者的 8 倍，极少出现隔 3 张牌插入 1 张牌的情况。洗多少次牌才能保证把牌的顺序完全打乱？

最近，概率论专家开始认真地研究这个问题，原因之一在于计算机被广泛地用于摹拟复杂的数学模型。美国统计学家帕西·达克尼斯是这方面的专家。帕西不但精通统计学，而且擅长魔术表演，此外他掌握了完美的弹式洗牌的技术。他和同事研究了洗牌的数学原理，发现了一个有趣的结论：无论用哪种方式洗牌，在最初几次洗牌之后，牌依然是有序的，但洗若干次之后，牌就会变成无序的。需要洗多少次牌才能从有序变成无序取决于具体的洗牌方式，而且需要的洗牌次数是可以计算的。利用最简单的洗牌方式（把最上面 1 张牌随机地插入其他 51 张牌中，不断重复）需要洗 240 次，利用标准的弹式洗牌 264（不是完美的弹式洗牌）需要洗 9 到 10 次。只洗 5 次牌显然太少了。

普通玩家洗牌通常可以保证牌的顺序随机，但是专家洗牌时，他们的高超技巧可能使牌的顺序有规律。理查德·爱普生说过一个笑话：真正随机的洗牌应当是在强风中抛出一副牌，然后让一个蒙上眼睛的醉鬼把牌捡回来。当然你不用这么费事。

分配奖金

亨利和斯坦利玩扑克赌钱，事先约定先赢 6 局者赢得全部赌注。进行 9 局以后，比分为 4—5，斯坦利领先，但此时斯

坦利有急事退出赌局，赌注应当如何分配?

在数学史上这是一个著名的问题。1654 年两个伟大的数学家——帕斯卡和费尔马通信讨论了这个问题，通常认为，关于概率论的系统化研究发端于此。帕斯卡和费尔马运用的方法不同，但他们得出了相同的结论。他们的基本思路是一致的：首先算出每一方为了最终获胜需要再赢多少局，以及完成这个目标的概率，然后根据概率值分配奖金。

在这个例子中，亨利最终获胜需要再赢 2 局，而斯坦利最终获胜需要再赢 1 局。假定在每 1 局中双方获胜的概率都是 1/2,则亨利完成目标的概率为 $1/2\times1/2=1/4$，所以亨利应当拿到全部赌注的 1/4，斯坦利应当拿到全部赌注的 3/4。分别以 x 和 y 表示斯坦利和亨利为了最终获胜需要再赢的局数，则他们完成目标的概率由 x 和 y 的值决定，用 p（x，y）表示。表 13.5 是计算结果。

表 13.5

	y = 1	y = 2	y = 3	y = 4
x = 1	1/2	3/4	7/8	15/16
x = 2	1/4	1/2	11/16	13/16
x = 3	1/8	5/16	1/2	21/32
x = 4	1/16	3/16	11/32	1/2

如果你想扩展表 13.5 以处理较大的 x 和 y 的值，也很简单。利用以下三个公式：

$$p(x+1,1)=p(x,1)/2$$

$$p(1,y)=1-P(y,1)$$

$$p(x+1, y+1) = (p(x+1, y) + p(x, y+1))/2$$

元音还是辅音

山姆和波利玩一个游戏。山姆翻开一本小说，在某一页中随机地挑出一个字母。波利的任务是猜这个字母是元音字母还是辅音字母。波利应当如何猜？他成功的概率是多少？

我们研究一下上面这段文字。在这段文字中出现了66个元音字母和105个辅音字母[①]，这证明了我们最初的直觉——辅音字母出现的频率高于元音字母。为了使成功的概率最大化，波利应当猜山姆挑出的字母是辅音字母。他的成功率就是那本小说中辅音字母出现的频率。

略微修改一下规则。山姆还是在书中随机地挑出一个字母，然后告诉波利这个字母之前的那个字母是不是元音字母。此时波利应当如何猜？山姆提供的额外信息是否有帮助？

我们还是研究一下本节的第一段文字。在这段文字中，元音字母后面跟一个辅音字母的情况出现了59次，而元音字母后面跟另一个元音字母的情况只出现了7次；另外，辅音字母后面跟一个元音字母的情况出现了59次，而辅音字母后面跟另一个辅音字母的情况只出现了45次。以这个统计资料为指导，我们的原则是：如果山姆告诉波利猜测对象之前是一个元音字母，波利就应当猜辅音字母；如果山姆告诉波利猜测对象之前是一个辅音字母，波利就应当猜元音字母。

这个问题其实是马尔可夫链的一个例子。1913年，俄国

① 作者指英文原文，把y当做元音字母——译者注

概率学家 A· A·马尔可夫分析了普希金的诗《叶夫根尼·奥涅金》中的2万个字母。他发现，元音字母出现的频率为43%，而辅音字母出现的频率为57%。所以，在没有额外信息的情况下，猜辅音字母是最佳选择。考虑相邻的两个字母的组合，马尔可夫发现元音字母后面跟一个辅音字母的频率为87.2%，而辅音字母后面跟一个元音字母的频率为66.3%。额外信息对作出决定是有帮助的。

266

马尔可夫把研究推进了一步。如果山姆告诉波利在猜测对象之前的两个字母是元音字母还是辅音字母，额外信息的帮助是否更大？与很多人的猜测相反，马尔可夫的结论是否定的。三个字母（n_1，n_2 和 n_3）相连，n_3 是猜测对象。山姆告诉波利 n_1 和 n_2 是元音字母还是辅音字母与仅仅告诉波利 n_2 是元音字母还是辅音字母没有差别。类似地，山姆告诉波利在猜测对象之前的三个（或更多）字母是元音字母还是辅音字母与仅仅告诉波利猜测对象之前的一个字母是元音字母还是辅音字母没有差别。

马尔可夫链被广泛应用于各个领域，如物理学、生物学、社会学、经济学、认识论、运动学等等。马尔可夫链的主要方法在于分析在一系列事件中一个事件发生的概率对其他事件发生的概率的影响。

马尔可夫链方法所依据的假设通常相当简单，分析过程也不复杂，当然有时候其模型是不完善的。在第八章中，我们讨论过斯诺克桌球赛和网球赛的例子，当时，我们假设每一局比赛的结果是互不影响的。当然，实际上某一局比赛的结果可能对其后的比赛结果有影响。比如说，一名选手输掉了第一局，此后他赢得第二局的概率可能会上升，因为在输掉第一局以后他会在第二局中更加卖力，而对手可能在赢得一局之后放松警

惕。当然也有相反的可能，一名选手在赢得一局之后赢得下一局的概率上升，即所谓的“从胜利走向胜利”。在具体的问题中，我们通常只考虑一个事件对紧随其后的一个事件的影响，而不考虑连续的两个事件对第三个事件的影响，因为我们不希望模型过于复杂。在我们正在讨论的例子中，我们就不考虑连续两个字母对第三个字母的影响。

在一系列事件中，我们希望通过已知的事件结果预测下一个将发生的事件的结果。如果我们可以假设一个事件的结果仅仅受前一个事件的影响，其他事件的结果可以忽略，则马尔可夫链方法就可以得到应用。其实这就是马尔可夫链方法的中心思想。在这个例子中，马尔可夫发现一个字母是元音或辅音对下一个字母是元音或辅音的影响可以表达如下：

	元音	辅音
元音	0.128	0.872
辅音	0.663	0.337

其中，0.128 表示一个元音字母后跟一个元音字母的概率，0.872 表示一个元音字母后跟一个辅音字母的概率，等等。

267

另一个例子是对气象资料的分析。某机关统计了 27 年来某地 12 月至 2 月的降雨记录，其中包括 2 437 天。把这些天分为“干型”和“湿型”两类，每一天要么属于“干型”，要么属于“湿型”。我们的目标是分析某一天的天气对第二天的影响。统计结果如下：

	干型	湿型
干型	0.750	0.250
湿型	0.338	0.662

也就是说，如果今天属于“干型”，则明天也属于“干型”的概率为 75%；而如果今天属于“湿型”，则明天属于“干型”的概率只有 33.8%。

在《叶夫根尼·奥涅金》中，三个字母 n_1、n_2 和 n_3 相连，已知 n_1 是元音字母，那么 n_3 是元音字母的概率是多少？用马尔可夫链可以回答这个问题。有两种情况使得 n_3 是元音字母：情况一，n_1、n_2 和 n_3 都是元音字母；情况二，n_1 和 n_3 是元音字母，而 n_2 是辅音字母。情况一对应的概率为 0.128×0.128，情况二对应的概率为 0.872×0.663，两个概率的和为 0.595。于是，答案为 59.5%。利用类似的方法可以求出，当 n_1 是元音字母时，n_3 是辅音字母的概率是 40.5%。

以上结论可以进一步推广。已知 n_1、n_2、……、n_i 是《叶夫根尼·奥涅金》中的 i 个字母，n_1 是元音字母，n_i 是元音字母的概率是多少？考虑 $i=4$ 的情况。有四种情况使得 n_4 是元音字母：1. 元音—元音—元音—元音；2. 元音—元音—辅音—元音；3. 元音—辅音—元音—元音；4. 元音—辅音—辅音—元音。分别计算四种情况对应的概率，求和后得出最后答案：0.345。利用类似的方法可以求出，当 n_1 是辅音字母时，n_4 是元音字母的概率是 0.498。随着 i 值的增加，初始信息对结论的影响减少。例如 $i=5$ 的情况：当 n_1 是元音字母时，n_5 是元音字母的概率是 41%；当 n_1 是辅音字母时，n_5 是元音字母的概率是 45%，n_1 的状态对 n_5 的状态的影响已经很小了。

当 $i=9$ 时，无论 n_1 是元音字母还是辅音字母，n_9 是元音字母的概率都是 43%。这个值就是元音字母在全文中出现的频率。

如果你想分析连续的多个事件对其后的事件的影响，马尔可夫链方法不适用。数学家开发了专门的算法对付这类问题。马尔可夫链的应用范围极其广泛，很多概率论方面的教科书细致地介绍这一工具。希望读者注意：当我们面对具体问题的时候，通常我们需要一个简化的数学模型，虽然我们的模型经常与实际情况有出入，但并不妨碍我们得出有意义的结论。 268

本福德法则

两个人（A 和 B）赌钱。A 随意选择一个真实的数。这个数可以是一条河的长度，一个人的月收入，一个镇的人口总数，等等。这个数是随意选择的，但必须是一个真实的数，不能是随机生成的：不能是计算机给出的随机数，也不能是从电话号码本或质数表中随机选择的数。当这个数确定以后，我们只关心这个数的第一位有效数字。（第一位有效数字指这个数中的第一个非 0 数字，例如 81，8.1 和 0.81 的第一位有效数字都是 8。）B 的任务是猜这个数的第一位有效数字是几。B 猜对的概率是多少？如果 A 投入的赌注为 10 英镑，B 应当投入多少赌注才能保证这个游戏对双方公平？

B 需要在 9 个数字中猜 1 个（因为第一位有效数字不可能是 0），如果他在 9 个数字中随机地猜 1 个，他猜对的概率为 1/9。当 A 投入的赌注额为 10 英镑时，B 投入的赌注额为 1.11 英镑可以保证游戏是公平的。然而，根据本福德法则可以得出一个惊人的结论：即使把 B 投入的赌注额增加一倍，B 依然可

以在游戏中获利。诀窍非常简单：B 每次都应当猜“1”。

如果 A 严格地按照规定选择数据，则数据的第一位有效数字是 1 的概率超过 30%。由于 A 每次投入 10 英镑的赌注，所以 B 平均每次赢得 3 英镑以上。我相信读者会非常困惑：既然数据是任意选择的，第一位有效数字是 1 的概率怎么会这么高？

这个发现归功于弗兰克·本福德，他在 1938 年写了《异常数字法则》一文。本福德无意中发现，对数表的前面通常被翻得很破旧，而后面的磨损通常很小。（由于计算机技术的进步和普及，当代人对于对数表已经很陌生了，但是在 25 年以前，对数表还是极其常用的数学工具。）这是怎么回事呢？本福德猜想，对数表的主人查阅以 1 或 2 开头的数据的频率比查阅以 [269] 8 或 9 开头的数据的频率高很多。其实在本福德以前，西蒙·纽克比在 1881 年就公布过同样的发现。但是本福德的研究引起了广泛的注意，所以大家把这一发现称为“本福德法则”。

本福德法则的完整表达为：任意选择一个真实的数，以十进制的形式表示，则第一位有效数字为 d 的概率为 $\log_{10}((d+1)/d)$。当 $d=1$ 时，对应的概率为 $\log_{10}2=0.301$，即第一位有效数字为 1 的概率为 30.1%；当 $d=2$ 时，对应的概率为 $\log_{10}(3/2)=0.176\,1$，即第一位有效数字为 2 的概率为 17.6%，等等。请参照表 13.6。

表中的第二行是根据本福德法则计算出的概率值。本福德本人用实际数据检验了以上结论，结果列在表中的第三行。本福德的实验数据来自于 20 个不同的领域，包含 2 万个实际数据。其数据来源包括：根据元素周期表得到的某些原子的重量，某些著名科学家的房子的数量，倒数表（一种数学用表）中的数据，等等。检验结果与理论计算的结果有一些出入，不

过随着检验的数据量的增加，出入将减少。表中的第四行是另一组检验结果，数据来源为《1980/1981AA 手册》中的 100 组相邻的条目。《1980/1981AA 手册》的内容是一些城镇的人口数，按照城镇的名称的字母顺序排列。表中的第五行是根据所罗门群岛某地的 1 243 个居民的月用电消费额统计的结果，来自于拉夫·雷米的一篇学术论文。所有这些检验都支持本福德法则。

这个结论难以置信，但得到了检验支持。我们必须设法解释这个现象。在以上讨论中我们用到一个意义含混的词：“任意选择一个真实的数”。选什么样的数才算是“任意选择一个真实的数”呢？在这个词的意义得到精确的界定之前，我们不可能对本福德法则给出严格的数学证明。所以下面仅仅给出粗略的描述。

我们经常应用两种类型的数据。一类数据与度量衡的单位无关，例如一个城镇的人口数。1 个人就是 1 个人，不涉及度量衡的单位；另一类数据则与度量衡的单位有关，例如一条河的长度。一条河的长度是多少取决于度量衡的单位，以米为单位的结论与以英尺为单位的结论不同，而且选择什么单位通常是任意的，可以是英里、英寸、公里、码等等。当一个数据在两个单位之间转换时，这个数据要乘以某个换算系数。例如，在英寸和厘米之间转换时，需要乘以 2.54，在码和英尺之间转换时，需要乘以 3。给定一组数据，统计出其中的第一位有效数字的分布频率，如果对每一个数据都乘以一个系数，则第一位有效数字的分布频率通常会改变。如果存在一组数据，对其中的每一个数据都乘以一个系数（系数不为 0），无论这个系数是多少，第一位有效数字的分布频率都保持不变，则分布频率必然符合本福德法则。

表 13.6

第一位有效数字	1	2	3	4	5	6	7	8	9
根据本福德法则计算的结果	30.1	17.6	12.5	9.7	7.9	6.7	5.8	5.1	4.6
本福德本人的检验结论	30.6	16.7	11.6	8.7	8.5	6.4	5.7	5.0	5.7
统计城镇人口的检验结论	28	20	13	11	5	10	9	3	1
统计居民用电量的检验结论	30.6	18.5	12.4	9.4	8.0	6.4	5.1	4.9	4.7

270

我们先做一个简单的试验。从一个初始数据出发，不断乘以一个固定的系数，然后统计其第一位有效数字的频率。以 6 为初始数据，以 2 为系数，则得到一组数据：6，12，24，48，96，192，……，其第一位有效数字分别为 6，1，2，4，9，1，……。把这组数据延长到第 100 位，统计第一位有效数字的频率，1～9 分别出现了 30，18，12，10，8，7，5，5，5 次。请注意，这个频率分布与本福德法则非常接近。把系数换成其他的数，如 1.5，4.2 等等，结论同样支持本福德法则。（当然，如果以 10 为系数，将遇到例外。）

当你"任意选择一个真实的数"的时候，你的选择范围其实符合本福德法则的分布规律，第一位有效数字为 1 的概率比为 9 的概率高很多。

以上文字算不上什么证明，因为我们的表述太不严格了。如果你想用实际数据检验本福德法则，精心选择数据来源可以得到相反的结论。例如，你可以选择在怀特岛上服刑的犯人的年龄作为数据来源。年龄不满 20 的犯人不会被送进怀特岛监狱，而 100 岁以上的犯人极其罕见，所以第一位有效数字为 1 的概率接近于 0，与本福德法则的结论相反。另外，如果我们

271

规定所有人口数在 200 到 999 之间的地区称为"村庄"，而后

以村庄的人口数为数据来源检验本福德法则，本福德法则肯定是不对的。已经有人搞出了本福德法则的数学证明，不过这些证明很难说是严格的，因为我们无法定义“任意选择一个真实的数”。如果你的目的就是和别人打赌，不必为此烦恼，每次都猜“1”，我保证你能赢钱。

把游戏换一个形式。A 随意选择一个真实的数，如果这个数的第一位有效数字不超过 3 则 B 获胜，否则 A 获胜。这个游戏对谁有利？非常简单，根据表 13.6，B 获胜的概率超过 60%。

哈尔·瓦林曾给出一个很有趣的建议：用本福德法则可以检验政府公布的财政数据是否真实。真实的数据应当符合本福德法则，如果我们发现结果与本福德法则有较大出入，就有理由怀疑有人做了手脚。

一个名叫马克·尼格林尼的会计师走得更远。他根据本福德法则的思想发展了一套方法，用来检验账目的真实性。如果有人对账本做过手脚，利用他的方法可以发现疑点。这种方法被称为“尼格林尼求和法”。

272

猜背面颜色

一个相当有趣的问题：一共有三张卡片（A、B、C），A 的两面都是红色，B 的两面都是黑色，而 C 有一面是红色，另一面是黑色。我从三张卡片中随机地抽出一张，放在桌面上，你只能看到卡片的一面，看不到背面。你的任务是猜背面的颜色。假设你看到的一面为红色，你将如何猜？

这个问题与第九章中的“挑盒子问题”非常相似，关键在

于如何判断额外信息对结论的影响。通讯理论权威华伦·韦沃在《幸运小姐》一书中讨论过这个问题。我相信很多人会这样分析：

> 由于这张卡片有一面是红色，所以这张卡片不可能是B。剩下的两种可能是A和C，这张卡片是A或C的概率相等，即1/2。如果此卡片是A，则背面为红色；如果此卡片是C，则背面为黑色。背面为红色的概率为1/2。
>
> 看起来很有道理，是不是？然而这个结论是错误的。事实上，背面为红色的概率为2/3。我与你打赌，如果背面为红色则我获胜，如果背面为黑色你获胜，赔率为1:1。如果你接受这个挑战，你将输掉很多钱。一战结束时，韦沃正在读研究生。当时他利用这个诡计赢了同学很多钱。当然，输钱的人不一定吃亏，他们学到了知识。

有很多方法证明背面是红色的概率为2/3。下面是两个证明：

> (1) 用a_1、a_2、b_1、b_2、c_1、c_2分别表示A、B、C的各个面，其中a_1、a_2和c_1是红色，b_1、b_2和c_2是黑色。最初，我们可以看到的那个面是6个面中的某1个的概率是相等的。当我们看到这个面是红色以后，这个面是b_1、b_2或c_2的可能性已被排除，这个面只能是a_1、a_2或c_1，每种可能的概率为1/3。如果这个面是a_1或a_2，则背面为红色；如果这个面是c_1，

则背面为黑色。因此，背面为红色的概率为 2/3。

(2) 假定你的策略是每次都猜背面的颜色与看到的这个面的颜色相同，这种策略的成功率为 p。这意味着两种情况：a. 当看到的这个面为红色时，猜背面为红色；b. 当看到的这个面为黑色时，猜背面为黑色。由于在这个问题中，红色和黑色是完全对称的，所以在两种情况下你的成功率相等，都是 p。(请读者仔细推敲。) 在三张卡片中，有两张两面的颜色相同。如果抽出的卡片为 A 或 B，这种策略成功；如果抽出的卡片为 C，这种策略失败。所以这种策略的成功率为 2/3。因此，当看到的这个面为红色时，猜背面为红色的成功率为 2/3。

273

相信你现在已经接受了我的观点：背面为红色的概率不是 1/2。但是，最初的分析错在哪里呢？注意这一句："这张卡片是 A 或 C 的概率相等"。确实，在我们没看到抽出的卡片的任何一面时，这个判断是正确的。然而，一旦我们发现这张卡片有一面为红色，额外信息就会对概率分布产生影响。我们看到的这一面为 a_1、a_2 或 c_1 的概率相等，即这一面是 a_1 或 a_2 的概率为这一面是 c_1 的概率的二倍。也就是说，这张卡片是 A 的概率为这张卡片是 C 的概率的二倍。

至少两张 A

在桥牌中，一手牌（13 张）至少包含两张 A 的概率仅为 26%。你和劳拉打赌，劳拉随机地取一手牌，如果其中包含至

少两张 A 则劳拉获胜，否则你获胜，赔率为 1:1。这个局面很简单，显然对你有利。然而，这个游戏可以衍生出两个复杂的问题：

问题一：劳拉随机地取一手牌，发现其中包含一张 A，此时她和你赌这手牌至少包含两张 A。劳拉的胜率是多少？

问题二：劳拉随机地取一手牌，发现其中包含黑桃 A，此时她和你赌这手牌至少包含两张 A。劳拉的胜率是多少？

这两种情况有差别吗？在问题二中，劳拉的胜率是否比在问题一中高？请不要轻率地下结论。

这个问题需要复杂的推算。52 张牌产生的变化太多，计算起来很麻烦，所以我们先考虑一种类似但简单得多的情况。假设一副牌只有 4 张：黑桃 A、K 和红桃 A、K。把一副牌分给两个人，每手牌两张。考虑三种情况：情况一，在没有任何额外信息的情况下，一手牌有 6 种可能，每种可能出现的概率相同，其中只有 1 种包含两张 A，因此一手牌包含两张 A 的概率为 1/6；情况二，劳拉拿到一手牌，发现其中包含一张 A，此时这手牌由两张 K 构成的可能已被排除，还剩下 5 种可能，每种可能出现的概率相同，因此这手牌包含两张 A 的概率为 1/5；情况三，劳拉拿到一手牌，发现其中包含黑桃 A，此时有三种可能已被排除，还剩下 3 种可能，这手牌包含两张 A 的概率为 1/3。显然，三种情况的结果各不相同。在一副牌包括 52 张的情况下，我们相信有类似的结论。

274 在进行完整的计算之前，我们先作一个粗略的估计。

在问题一中，最初劳拉手中的 A 的张数可能是 0、1、2、3 或 4，对应的概率分别为 p_0、p_1、p_2、p_3、p_4。在劳拉发现她手中有一张 A 之后，p_0 已被排除，所以她手中至少有两张 A 的概率为 $(p_2+p_3+p_4)/(p_1+p_2+p_3+p_4)$。由于一手牌包含

3或4张A的情况很少见，所以p_3和p_4很小，可以忽略。另外，平均每手牌包含1张A，所以直观的结论是p_1大于p_2。因此，问题一的答案在1/2附近。

在问题二中，劳拉的牌由1张黑桃A和另外12张牌构成，这12张牌是从51张牌（其中包括除黑桃A以外的3张A）中随机选择的。劳拉手中有至少两张A即这12张牌中至少有一张A。51张牌中每张牌是A的概率为3/51，在12张牌中平均包括12（3/51）张A，约等于2/3。这12张牌中A的张数可能是0、1、2或3，张数是0或1的概率极高，而且张数是1的概率高于张数是0的概率。因此，劳拉手中至少有两张A的概率在1/2～2/3之间。

现在进行严格的计算。在问题一中，我们分别计算出p_0和p_1的值。劳拉手中至少有一张A的概率为$1-p_0$，至少有两张A的概率为$1-p_0-p_1$。在劳拉发现她手中有一张A的前提下，她手中有至少两张A的概率为$(1-p_0-p_1)/(1-p_0)$，约等于37%。

在问题二中，计算更简单，我们只要算出在12张牌中没有A的概率就可以了。在51张牌中取12张，有${}^{51}C_{12}$种可能，其中${}^{48}C_{12}$种不包含A，所以51张牌中至少有1张A的概率为$1-{}^{48}C_{12}/{}^{51}C_{12}$种约等于56%。这也就是问题二的答案。

我相信，很多读者在仔细研究以上讨论之前认为问题一和问题二的答案是一样的。他们可能这样认为："当劳拉发现手中有一张A时，这张A是黑桃的概率与是红桃（或方块、梅花）的概率是一样的，而确定这张A是否为黑桃并未提供更多的有效信息。"事实上，这张A是否为黑桃至关重要。无论这张A是否为黑桃，都将排除若干种可能性，而在确定其为黑桃的情况下，被排除的可能性更多。请仔细研究简化后的例

子（一副牌只有 4 张的情况）。

在这个问题中，很多人坚持错误的结论。他们坚信，两种情况提供的信息是相同的。我承认，说服他们是很困难的，因为我们无法把一手牌的所有可能情况都列出来。

这个问题与第九章中的“挑盒子问题”和上一节中的“猜背面颜色问题”类似。我希望我的论证足以证明以上分析是正确的，但是这个结论却受到专家的质疑。我们最初考虑的 4 张牌构成一副牌的例子是马丁·格德纳给出的，此人是《科学美国人》的专栏作家，负责“趣味数学”专栏。格德纳的结论是，当已知手中有一张 A 时，手中至少有两张 A 的概率为 1/5；而当已知手中有黑桃 A 时，手中至少有两张 A 的概率为 1/3。华伦·韦沃认为格德纳的分析有误。他的观点是，最初 6 种可能出现的概率相等，而当其中的一种（或几种）可能被排除之后，剩下的几种可能出现的概率已经不相等了——与“猜背面颜色问题”中的情况一样。

我很欣赏华伦·韦沃对这两个问题所作的类比，然而两种情况有本质的差别。在“猜背面颜色问题”中，发现卡片的一面是红色排除了某些可能，而且在剩下的可能中，最初每种可能出现的概率相等，而获得额外信息之后概率不再相等。在扑克牌的问题中，并不存在类似的情况。

韦沃认为，这个问题中包含很多复杂的因素我们尚未涉及。我接受这个观点。假设劳拉拿到一手牌，然后发现她的牌符合以下 5 种情况：

○ 有 1 张 A；

○ 有黑桃 A；

○ 没有草花；

○ 有黑桃；

○ 至少有 5 张比“9”大的牌。

但是，劳拉并不完全告诉我们这些信息，而是随意地告诉我们其中的 1 条（或几条）。此时，问题变得极其复杂，我们需要分析很多因素。

为什么韦沃的观点与我们不同？我猜想韦沃考虑了这些复杂的因素。韦沃采用的模型与我们不同。在第一章中我说过，人类拥有概率的直觉。当我们面对一个经验问题时，首先把问题转化为一个模型，然后才根据模型计算概率。不同的人掌握的信息量不同，他们可能针对同一个问题构造出不同的模型，每一个模型都是有道理的，但结论不同。

对哪一方有利?

当一场比赛的规则确定之后，通常可以断定比赛对哪一方有利。然而，什么叫做“对某一方有利”，有时候我们并没有确切的定义。托马斯·卡沃用一系列例子说明这个问题。这些例子表明，比赛的长度是一个至关重要的因素。

假设 A 和 B 比赛，比赛进行若干局。在每局中 A 得 2 分，而 B 要么得 3 分，要么得 0 分。B 得 3 分的概率 p 是固定的，$p=0.618$。比赛的胜负由双方的累积得分决定，不考虑每方获胜的局数。这个比赛对谁有利？假设比赛只进行 1 局，A 的得分固定为 2，而 B 有接近 62%的可能得 3 分，所以 B 胜出的概率较高。然而，假设比赛进行两局，则对 A 有利。A 的得分为 4，而 B 获胜的条件是每局都得 3 分，概率为 p^2，接近 38%。当比赛进行的局数更多时，A 获胜的概率较高。

调整一下得分规则，将得到一个有趣的例子。假设在每局

277 中 A 得 500 分，而 B 以 99%的概率得 501 分，以 1%的概率得 0 分。如果比赛仅进行 1 局，B 获胜的概率为 99%。而如果比赛进行 500 局，A 的得分为 2.5 万，B 的获胜条件是每局都得 501 分。只要 B 有 1 次得 0 分，他的累积得分就不可能高于 24 999。B 获胜的概率为 $(0.99)^{500}$，约等于 0.006 6。当比赛局数从 1 变到 500 时，B 获胜的概率从 99%降到 1%以下，差别巨大。如果我们进一步调整得分规则，可以得到更惊人的数据变化。

卡沃构造了一套复杂的比赛规则。在这种规则下，B 获胜的概率随着比赛局数的增加上下震荡。最初，B 获胜的概率接近于 1，而若干局之后，B 获胜的概率接近于 0；再进行若干局，B 获胜的概率又接近于 1，而且距离 1 更近；再进行若干局，B 获胜的概率又接近于 0，而且距离 0 更近；以此循环。比赛对哪一方有利取决于进行多少局。结果可能是这样：如果比赛进行 500，1 000 或 3 000 局，B 获胜的概率接近于 1；而如果比赛进行 800 或 2 000 局，B 获胜的概率接近于 0。构造这样一种规则不是一件容易的事，但这种规则确实存在。

我们总说，长期比赛的结果对强者有利。然而，谁是强者取决于比赛进行多长。《伊索寓言》支持这个判断。在龟兔赛跑中，兔子是否比乌龟强呢？ 278

附录 1　计数

在扑克牌、彩票和轮盘赌等例子中，我们通常假设每一种结果出现的概率是相同的。在这类问题中，解决问题的关键往往在于判断一共有多少种可能出现的结果，计数成为核心技术。在计数过程中，乘法原理极为常用：

如果完成事件 A 有 m 种方法，而在完成 A 以后完成事件 B 有 n 种方法，则完成 A 和 B 有 $m\times n$ 种方法。

例如，把三个字母 A、B、C 排序，一共有多少种排序方法？为了解决这个问题，我们可以排列出所有可能：

ABC　ACB　BAC　BCA　CAB　CBA

于是，答案为 6。不过，如果把问题变成为五个字母 A、B、C、D、E 排序，这种笨方法就不大管用了。这个问题可以用乘法原理处理：

> 选择排在第一位的字母有 $m=5$ 种方法。在第一位字母确定以后，选择排在第二位的字母有 $n=4$ 种方法。因此，选择前两位字母有 $5\times4=20$ 种方法。在前两位字母确定以后，选择排在第三位的字母有 3

种方法。于是，选择前三位字母有 $20\times3=60$ 种方法。此后，选择排在第四位的字母有 2 种方法。于是，选择前四位字母有 $60\times2=120$ 种方法。在前四位字母确定以后，选择排在最后的字母只有 1 种方法。所以，一共有 $120\times1=120$ 种方法为五个字母排序。

请注意以上计算过程。计算结果是连续的五个数的乘积：

$$5\times4\times3\times2\times1=120$$

为书写方便，我们用 5！表示 $5\times4\times3\times2\times1$，读作“5 的阶乘”。当字母个数为 3 时，排序的方法有 $3!=3\times2\times1=6$ 种。一般性的结论如下：

结论一：

为 K 个对象排序，排序方法有

$$K!=K\times(K-1)\times(K-2)\times\cdots\times3\times2\times1 \text{ 种}$$

随着 K 值的增加，K！的值迅速上升。为一副扑克牌排序，排序方法有 52！种，这个数长达 68 位。表 A1.1 是 K！的几个值。

表 A1.1

K的值	1	2	3	4	5	6	7	8
K！的值	1	2	6	24	120	720	5040	40320

另一个问题：从 6 个对象中选出 2 个，有多少种选法？

从 6 个对象中选出第一个对象，有 $M=6$ 种选法；然后在

剩下的5个对象中选出第二个对象，有 $N=5$ 种选法。所以，一共有 $6\times5=30$ 种选法。但是，每一种选出的结果都被计算了两次，例如对于｛A，B｝而言，AB和BA是同一种结果，但被分别计算。所以，最后的结论是30除以2，即从6个对象中选出2个，不考虑对象的次序，有15种选法。

一般地，从K个对象中选出2个，第一个对象有K种选法，第二个对象有 $K-1$ 种选法，一共有 $K(K-1)$ 种选法。由于每一种结果被计算了两次，这个值除以2就是结论。

结论二：

从K个对象中选出2个，不考虑对象的次序，一共有

$$K\times(K-1)/2 \text{ 种选法}$$

如果把问题变成从K个对象中选出3个，不考虑对象的次序，一共有多少种选法呢？解法是类似的。第一步，算出在考虑对象的次序的情况下有多少种选法；第二步，计算每一种结果被计算了多少次。以 $K=7$ 的情况为例：在考虑对象次序的情况下，一共有 $7\times6\times5=210$ 种选法；每一种结果被计算了 280

$$3\times2\times1=6$$

次。最后的结论是 $210/6=35$。

以上计算过程即 $\frac{7\times6\times5}{3\times2\times1}=35$

类似地，当 $K=10$ 时，结论为 $\frac{10\times9\times8}{3\times2\times1}=120$；当 $K=12$ 时，结论为 $\frac{12\times11\times10\times9\times8}{5\times4\times3\times2\times1}=792$。在概率论中，我们经常用到类似的算法。我们用符号 ${}^{12}C_5$ 表示在12个对象中选出5个（不考虑次序）的选法种数，读作“12选5”。以上几个结果可以表示为 ${}^{10}C_3=120$，${}^{7}C_3=35$，等等。结论二可以表示为 ${}^{K}C_2$

$=K(K-1)/2$。类似地有

$${}^{K}C_3=\frac{K\times(K-1)\times(K-2)}{3\times2\times1},$$

$${}^{K}C_4=\frac{K\times(K-1)\times(K-2)\times(K-3)}{4\times3\times2\times1}，等等$$

结论三：

从 K 个对象中选出 r 个，不考虑对象的次序，一共有 ${}^{K}C_r$ 种选法。

我们经常遇到这样的情况：我们想知道一个事件有多少种可能结果，但是我们不可能把各种可能一一列出，因为数据量过于庞大。例如，打桥牌时，从 52 张牌中选出 13 张构成一手牌，一共有 635 013 559 600 种可能结果；在买大英帝国彩票时，从 49 个号码中选出 6 个，一共有 13 983 816 种可能结果。此时，这个计算公式能帮大忙。

表 A1.2：帕斯卡三角形的一部分

1

1 1

1 2 1

1 3 3 1

1 4 6 4 1

1 5 10 10 5 1

1 6 15 20 15 6 1

1 7 21 35 35 21 7 1

1 8 28 56 70 56 28 8 1

我们经常把${}^{K}C_r$的值按表 A1.2 的形式排列，这个表称为"帕斯卡三角形"。从上向下，每一行分别表示 K = 0，1，2……的情况；从左向右，斜排的每一行分别表示 r = 0，1，2……的情况。表中的第 K + 1 行的第 r + 1 个值就是${}^{K}C_r$的值。 281

例如，表中的第 7 行为 1，6，15，20，15，6，1，这 7 个数分别是${}^{6}C_0$, ${}^{6}C_1$, ${}^{6}C_2$，……${}^{6}C_6$的值。这个表有三个特征：1. 左右对称；2. 两条边上的值为 1；3. 如果一个值不在边上，则这个值等于其左上角上的值与其右上角上的值的和。这些特征通常可以简化运算。

下面考虑另一个问题。抛 5 次硬币，有多少种可能结果？根据乘法原理，答案为 32。抛第 1 次可能有"正"、"反"2 种结果，抛第 2 次也可能有"正"、"反"2 种结果，所以抛前 2 次可能有"正正"、"正反"、"反正"、"反反"4 种可能结果 $(2 \times 2 = 4)$；同样，抛前 3 次可能有 $4 \times 2 = 8$ 种结果，抛前 4 次可能有 $8 \times 2 = 16$ 种结果，而抛 5 次可能有 $16 \times 2 = 32$ 种结果。 282

结论四：

抛 n 次硬币，可能有 2^n 种结果。

这个结论可以推广到掷骰子的情况中。掷 1 次骰子可能有 6 种结果，掷 2 次骰子可能有 $6 \times 6 = 36$ 种结果，掷 3 次骰子可能有 $6 \times 6 \times 6 = 216$ 种结果，等等。

结论五：

一个事件可能有 K 种结果，把这个事件重复 n 次，则有 K^n 种可能结果。

随着 n 的值的增加，2^n 和 K^n 的值迅速增加。连续抛 20 次

硬币，可能结果超过100万种。连续掷8次骰子，可能结果也将超过100万种。此时，我们已经没有能力把各种可能结果一一罗列出来了。

现在我们利用以上五个结论处理一个实际问题。

在第十章中我们讨论过赌场扑克，当时我介绍了各类牌（同花顺、四条、对子等等）有多少种。那些数字是如何计算出来的？

在52张牌中抽出5张，一共有

$$^{52}C_5 = \frac{52\times51\times50\times49\times48}{5\times4\times3\times2\times1} = 2\ 598\ 960 \text{ 种可能}$$

其中有多少种是“对子”？1副对子包括两张大小相同的牌，其大小有13种可能（A，K，……2），每1种大小又涉及4种花色。在4种花色中选择两种，有$^4C_2=6$种选法。所以，组成对子的两张牌可以有$13\times6=78$种。在这两张牌确定以后，再加上另外3张牌即可以构成1副对子。这3张牌的大小各不相同，并且与组成对子的那两张牌大小也不同，所以确定这3张牌的大小有$^{12}C_3=220$种可能。在这3张牌的大小确定以后，每张牌的花色可以是4种花色中的任意1种，所以确定这3张牌有$220\times4\times4\times4=14\ 080$种可能。最后，构成对子的牌一共有$78\times14\ 080=1\ 098\ 240$种。

另外，有多少副牌是“两对”？在1副两对中，有4张牌组成两组对子。确定两组对子的大小有$^{13}C_2=78$种可能。此外，每组对子的颜色有$^4C_2=6$种可能。所以，这4张牌的组成有$78\times6\times6=2\ 808$种可能。在这4张牌确定以后，第5张牌可以是$11\times4=44$张牌中的1张，所以，构成两对的牌一共有$2\ 808\times44=123\ 552$种。

有多少副牌是“三条”？1副三条包括3张大小相同的牌，

其大小有 13 种可能。在大小确定以后，这 3 张牌的花色有$^4C_3$ $=4$ 种可能。于是，这 3 张牌的组成有 $13\times4=52$ 种可能。在这 3 张牌确定以后，考虑另外两张牌。这两张牌的大小有$^{12}C_2$ $=66$ 种可能，每张牌的花色有 4 种可能，所以，构成三条的牌一共有 $52\times66\times4\times4=54\ 912$ 种。 283

有多少副牌是“三带二”？1 副三带二包括 3 张大小相同的牌和另外 2 张大小相同的牌。与三条的情况一样，3 张相同的牌的组成有 $13\times4=52$ 种可能。在这 3 张牌确定以后，考虑另外两张牌。这两张牌的大小有 12 种可能，花色有$^4C_2=6$ 种可能，所以，构成三条的牌一共有 $52\times12\times6=3\ 744$ 种。

有多少副牌是“四条”？1 副四条包括 4 张大小相同的牌，这 4 张牌的组成有 13 种可能。剩下的 1 张牌可以是其余的 48 张牌中的任意 1 张，构成四条的牌一共有 $13\times48=624$ 种。

在以上各种情况中，都包括两张（或更多）大小相同的牌。而以下几种情况则不同，其特征为每张牌的大小都不一样。首先考虑“顺子”和“同花顺”。以牌的大小为标准，顺子可以分为 10 种：从 A2345 到 10JQKA。在牌的大小确定以后，5 张牌中的每 1 张的花色有 4 种可能，所以花色的组合有 $4\times4\times4\times4\times4=1\ 024$ 种。在这 1 024 种花色组合种，有 4 种组合使得 5 张牌的花色完全相同。因此，同花顺有 $10\times4=40$ 种，而顺子有 $10\times(1\ 024-4)=10\ 200$ 种。

在 1 种花色的 13 张牌种选出 5 张，即可构成 1 副“同花”，一共有$^{13}C_5=1\ 287$ 种。其中有 10 种是同花顺，所以每一种花色中的同花有 1 277 种。一共有 $1\ 277\times4=5\ 108$ 种同花。

现在考虑有多少副牌属于“AK”。先考虑 5 张牌的大小。在 5 张牌中，有 1 张 A 和 1 张 A，另外 3 张牌大小各不相同，而且不是 A 和 K，所以牌的大小共有$^{11}C_3=165$ 种可能。如果

这3张牌恰好是“QJ10”，则整副牌成为顺子，所以要排除这种可能，即牌的大小有164种可能。在牌的大小确定以后，5张牌的花色组合有1 020种可能，所以AK一共有164×1 020=167 820种。

可以用两种方法计算“烂牌”有多少种。由于1副牌总共有$^{52}C_5$种可能，以上我们已经算出了所有不是烂牌的牌有多少种，做一次减法就可以知道有多少种烂牌。下面介绍第二种计算方法。

首先考虑牌的大小。烂牌可以分为三类：第一类包含一张A但不包含K；第二类相反，包含一张K但不包含A；第三类既不包含A也不包含K。在第一类牌中，除A以外的4张牌的大小各不相同，并且不是A和K，所以一共有$^{11}C_4=330$种可能。如果这4张牌恰好是2、3、4、5，则整副牌成为顺子，所以应当排除这种可能，即牌的大小有229种可能。类似地，第二类牌的大小也有229种可能（计算方法相同）。在第三类牌中，5张牌都不是A和K，而且大小各不相同，所以共有$^{11}C_5=462$种可能。在这462种可能中，有7种是顺子（23456，34567，……，8910JQ），所以第三类牌的大小有455种可能。把三个值加起来，得到烂牌的大小一共有329+329+455=1 113种可能。在牌的大小确定以后，确定5张牌的花色有1 020种可能，所以烂牌一共有1 113×1020=1 135 260种。

你可以比较两种计算方法的结果是否一致。如果结果一致，就说明我们的计算是正确的。

附录2　概率

如果一个事件的各种可能结果出现的机会相同，则只要我们算出一共有多少种结果，立刻就可以得出每种结果出现的概率。例如，在掷骰子的例子中，每种结果出现的概率为1/6，在轮盘赌的例子中，每种结果出现的概率为1/37，等等。

结论一：

如果一个事件有N种可能结果，每种结果出现的机会相同，则每种结果出现的概率为1/N。

在数学上，这种情况（指每种结果出现的机会相同）被称为“均匀分布”。通常，我们把N种结果表示为{1，2，3，……，N}。在均匀分布中，每种结果出现的概率相同，所以问题的关键在于发现一共有多少种可能结果，而解决这个问题的技巧我们已经掌握了。

有时候，我们把若干种结果作为一类结果，研究这一类结果出现的概率。例如，掷一次骰子，有时候我们关心结果属于{2，4，6}的概率（即出现的点数为偶数的概率），有时候我们关心结果属于{1，2，3，4}的概率（即出现的点数不超过

4 的概率)。当我们使用“结果”这个词时，有时候指一种结果，有时候指某一类结果。

结论二：

在一个均匀分布中有 N 种可能结果，有一类结果恰好包含其中的 K 种，则这类结果出现的概率为 K/N。

以抛硬币为例。在附录 1 中，我们知道抛 5 次硬币可能得到 $2^5=32$ 种结果。恰好出现 3 次正面的概率是多少?

在 5 次中选出 3 次结果为正面，一共有 ${}^5C_3=10$ 种可能，所以答案为 10/32。

在桥牌中，一手牌恰好包含 4 张红心的概率是多少? 根据附录 1，一手牌一共有 ${}^{52}C_{13}$ 种可能。一手牌恰好包含 4 张红心，意味着在 13 张红心中选出 4 张（${}^{13}C_4=715$），在 39 张非红心中选出 9 张（${}^{39}C_9=211\ 915\ 132$）构成一手牌。符合条件的牌恰好有 ${}^{13}C_4\times{}^{39}C_9$ 种，所以答案是

285

$$\frac{{}^{13}C_4\times{}^{39}C_9}{{}^{52}C_{13}}\approx 0.238\ 6$$

一手牌恰好包含 9 张红心的概率是多少? 根据同样的方法可以得出结论：

$$\frac{{}^{13}C_4\times{}^{39}C_4}{{}^{52}C_{13}}\approx 0.0009\ 26$$

这个概率低于万分之一。

把一副牌洗乱，随即抽出一张，每一张牌被抽出的概率都是 1/52。于是

○ 抽出一张黑桃的概率为 13/52;

○ 抽出一张草花的概率为 13/52。

所以，抽出一张黑牌的概率 = 抽出一张黑桃或草花的概率

= 13/52 + 13/52 = 1/2。另外一个例子与此不同：

○ 抽出一张黑桃的概率为 13/52；

○ 抽出一张 K 的概率为 4/52；

○ 抽出一张黑桃或抽出一张 K 的概率为 16/52。

最后一个值不是前两个值的和。如果你不小心，你可能得出 17/52 的错误结论。关键在于有一张牌（黑桃 K）既是黑桃又是 K，这张牌被计算了两次。在一般情况下，当我们计算“结果为 A 或 B 的概率”时，需要注意有多少种结果使得 A 和 B 重合。如果 A 表示“抽出一张黑桃”，而 B 表示“抽出一张草花”，则两种结果不重合。这种情况成为“排斥事件”。

结论三：

分别以 a、b 表示事件 A、B 发生的概率，以 c 表示 A 或 B 发生的概率。如果 A 和 B 是排斥事件（即 A 和 B 不能同时发生），则 $c=a+b$；如果 A 和 B 不是排斥事件，则 $c<a+b$。（即使 A 和 B 不属于均匀分布，这个结论依然成立。） 286

刚才我们已经得出一个结论：在桥牌中，一手牌恰好包括 4 张红心的概率为 0.238 6。同理，一手牌恰好包括 4 张黑桃的概率也是 0.238 6。那么，一手牌包括 4 张红心或包括 4 张黑桃的概率是多少？由于“一手牌包括 4 张红心”和“一手牌包括 4 张黑桃”这两件事不是排斥事件（因为一手牌可以既包括 4 张红心又包括 4 张黑桃），所以这个概率低于 0.238 6 + 0.238 6。相反的情况：“一手牌包括 9 张红心”和“一手牌包括 9 张黑桃”这两件事是排斥事件（因为一手牌只有 13 张，不可能既包括 9 张红心又包括 9 张黑桃）。进一步说，“一手牌包括 9 张红心”、“一手牌包括 9 张黑桃”、“一手牌包括 9 张方块”、“一手牌包括 9 张草花”这四件事是排斥事件，所以一手牌中

有 1 个花色包括 9 张牌的概率为 4×0.000 092 6 = 0.000 37。平均每 2700 手牌中会出现 1 次这种情况。

在第一章中我们介绍过“独立事件”这个概念，现在我们详细地解释这个概念。在 1 副扑克牌中随机地抽出 1 张，则

○ 这张牌是黑桃的概率为 13/52 = 1/4；

○ 这张牌是 K 的概率为 4/52 = 1/13。

如果判断这两个事件是否独立？如果这两个事件同时发生的概率等于两个事件各自的概率的乘积，则这两个事件是独立事件。在这个例子中，抽出的牌既是黑桃又是 K 的概率为 1/52（即抽出的牌为黑桃 K），而 1/52 = 1/4 × 1/13，所以这两个事件是独立的。

由于每种花色中都有且仅有 1 张 K，所以在我们确知 1 张牌是黑桃以后这张牌是不是 K 的概率不受影响。同样，在我们确知 1 张牌是 K 以后这个信息对于判断其是不是黑桃也没有帮助。这就是独立性的含义。面对具体问题时，通常我们可以轻易地判断出两个事件是否独立。在掷骰子或抛硬币时，连续掷（或抛）两次，我们通常认为两次的结果是相互独立的；而在一副扑克牌中抽出 1 张牌，连续抽两次，第一次的结果对第二次的结果有影响。

同时掷 4 枚骰子，要么其中至少出现 1 个 6 点，要么 1 个 6 点也没有。哪一种情况的概率更大？事件 A 是“至少出现 1 个 6 点”（概率为 a），事件 B 是“1 个 6 点也没有”（概率为 b），显然 a + b = 1。如果我们可以算出 b 的值，利用 1 − b 就可以算出 a 的值。

287

1 枚骰子的结果是 6 点的概率为 1/6，不是 6 点的概率为 5/6。由于 4 枚骰子的结果是相互对立的，所以 4 枚骰子的结果都不是 6 点的概率为 $(5/6)^4$，即 b = $(5/6)^4$ = 625/1 296，a

$=1-625/1\ 296=671/1\ 296$。由于 $671>625$，所以“至少出现1个6点”的概率较大。

当然你可以有另一种方法计算 a 的值。分别计算得到的6点的个数为1、2、3以及4的概率，而后求4个概率值的和。这种方法复杂得多。而利用 $a+b=1$ 这个关系，可以非常便捷地得出结论。

这种方法很常用。一个事件要么发生，要么不发生，所以两个概率值之和必为1，知道了其中的一个，另一个也就知道了。

非均匀分布

一个事件有若干种结果，在许多场合下，各种结果出现的机会不相等。比如说掷一次骰子，结果要么是6，要么不是6，概率分别为1/6和5/6。考虑两个问题：掷10次骰子，至少出现4个6点的概率是多少？连续掷骰子，第一个6点出现在第10次或第10次以后的概率是多少？数学家发明了一种处理这类问题的一般方法，即“伯努利试验”。伯努利试验要求三个条件：

(1) 每次试验有两种可能结果。分别用S和F表示两种结果；

(2) 每次试验的结果是相互独立的；

(3) 在每次试验中，结果为S（或F）的概率相同。

连续进行若干次试验，假定在第n次试验中首次得到S。用P(1)，P(2)，P(3)……分别表示 $n=1$，2，3……的概率。这些概率值分别为多少？

首先考虑 $n=5$ 的情况。在第5次试验中首次得到S就意味着前4次试验的结果都是F而第5次试验的结果为S，5次试验的结果依次为FFFFS。用p表示在1次试验中得到结果S的概率，则得到结果F的概率为 $1-p$。由于每次试验的结果是相互独立的，所以5次试验的结果依次为FFFFS的概率为

288

$$(1-p)\times(1-p)\times(1-p)\times(1-p)\times p=(1-p)^4p$$

用同样的方法可以得出当 $n=8$ 时的结论：在第8次试验中首次得到S的概率为 $(1-p)^7p$。

结论四：

在一系列伯努利试验中，首次S出现在第n次试验中的概率为 $(1-p)^{n-1}p$。(数学上称为"几何分布"。)

以抛硬币为例，S表示结果为正面。此时 $p=1-p=1/2$。当 $n=1，2，3$……时，对应的概率值分别为1/2，1/4，1/8……。这就是第四章中讨论的圣彼得堡游戏的结论。

随着n的值的增加，$(1-p)^{n-1}p$ 的值递减，因为n每增加1，$(1-p)^{n-1}p$ 的值就要乘以 $(1-p)$，而 $1-p$ 的值小于1。这意味着，首次S出现在第1次试验中的概率高于首次S出现在第2次试验中的概率，而首次S出现在第2次试验中的概率高于首次S出现在第3次试验中的概率，如此等等。

如何计算首次S出现在第12次或第12次以后的概率？有两种明显的方法。其一是算出 $n=12，13，14$……时对应的概率值，而后求和。更简单的方法是求出前11次的结果都是F的概率，因为前11次的结果都是F等同于首次S出现在第12次或第12次以后。结论是 $(1-p)^{11}$。

结论五：

在一系列伯努利试验中，首次 S 出现在第 n 次试验中或第 n 次试验以后的概率为 $(1-p)^{n-1}$。

除了 S 的首次出现以外，我们经常关心另一个问题：在连续 n 次伯努利试验中，恰好出现 r 次 S 的概率是多少？例如，连续掷 5 次骰子，恰好出现 2 次 6 点的概率是多少？这个问题可以分两步解决。

第一步，考虑有多少种结果符合条件（即恰好出现 r 次 S)。在掷骰子的例子中，这样的结果是 SSFFF，SFSFF，SFFSF，等等，其特征是在 5 次试验中有 2 个 S 和 3 个 F，一共有 $^5C_2$ 种。 289

第二步，考虑每种结果出现的概率。SSFFF，SFSFF，SFFSF 这几种情况对应的概率分别为

$$p\times p\times(1-p)\times(1-p)\times(1-p),$$
$$p\times(1-p)\times p\times(1-p)\times(1-p),$$
$$p\times(1-p)\times(1-p)\times p\times(1-p),$$

等等。这些概率都等于 $p^2(1-p)^3$，因为每个表达式中都恰好包含 2 个 p 和 3 个 $(1-p)$。把这些值加起来，就得出结论。由于这些值一共有 $^5C_2=10$ 个，所以结论是 $10p^2(1-p)^3$。

以上结论可以推广到一般情况。

结论六：

在连续 n 次伯努利试验中，恰好出现 r 次 S 的概率是

$$ {}^{n}C_{r}p^{r}\ (1-p)^{n-r}。 $$

在数学上，这种情况被称为“二项分布”。

在二项分布中，r的值可以是0～n的任意整数。随着r的值的增加，最初概率值递增，在达到最大值之后逐步递减。(最大值出现在什么位置取决于p的大小，当p较大时最大值的位置靠前，当p较小时最大值的位置靠后。)

下面是二项分布的几个例子。

(1) 5次硬币，恰好出现3次正面的概率是多少?

问题符合伯努利试验的条件。分别以S和F表示正面和反面，$p=1-p=1/2$，$n=5$，$r=3$，根据二项分布得出结论：

$$ {}^{5}C_{3}\ (1/2)^{3}\ (1/2)^{2}=10\ (1/2)^{5}=10/32 $$

(2) 在双陆棋中，一个玩家掷4次骰子(每次掷两枚)，恰好两次出现两枚骰子的点数相同的情况的概率是多少?

同时掷两枚骰子，两枚骰子的点数相同的概率为1/6(请读者自行证明)。于是$p=1/6, 1-p=5/6$。答案为

$$ {}^{4}C_{2}\ (1/6)^{2}\ (5/6)^{2}=6\ (1/36)\ (25/36)\ =150/1\,296 $$

相似的问题：一个玩家掷4次骰子(每次掷两枚)，至少有两次出现两枚骰子的点数相同的情况的概率是多少?

我们需要考虑三种情况：两枚骰子的点数相同的情况恰好出现2、3、4次。第一种情况的概率就是150/1296，第二种情况的概率是

$$ {}^{4}C_{3}\ (1/6)^{3}\ (5/6)^{1}=20/1\,296 $$

第三种情况的概率是

$$ {}^{4}C_{4}\ (1/6)^{4}\ (5/6)^{0}=1/1\,296 $$

答案是

$(150/1\ 296)+(20/1\ 296)+(1/1\ 296)=171/1\ 296=19/144$

(3) 一名推销员每次推销的成功率为 1/10。此人进行 10 次推销，至少成功 2 次的概率是多少？

假定每次推销的成果是相互独立的，则问题符合二项分布。在 10 次推销中恰好成功 r 次的概率为

$$^{10}C_r\ (1/10)^r\ (9/10)^{10-r}$$

分别计算当 r = 2，3，4，……，10 时的概率，而后求和就得出答案。这样计算相当麻烦，所以我们找一种偷懒的算法。我们可以先求出成功的次数不足 2 次的概率，然后用 1 减去这个值就得出结论。成功的次数不足 2 次意味着 r = 0 或 1，而 [291]

$$^{10}C_0\ (1/10)^0\ (9/10)^{10}=\ (9/10)^{10}$$

$$^{10}C_1\ (1/10)^1\ (9/10)^9=\ (9/10)^9$$

这两个值的和为 0.736 1。所以答案是 1 - 0.736 1 = 0.263 9，接近 26%。从两个概率值的和为 1 出发通常可以简化运算过程。数学家非常喜欢偷懒。

(4) 在第八章中提出的问题：两名选手对抗，一名选手在每局中的胜率如何影响他赢得整场比赛的概率？

假定每局比赛的结果是相互独立的。用 1/2 + y 表示选手 A 赢得 1 局比赛的概率。整场比赛进行 2N + 1 局时，A 获胜的条件是赢得 N + 1 局以上。分别计算 r 不低于 N + 1 时对应的概率，而后求和就得到 A 赢得整场比赛的概率。

当比赛为三局两胜（即 2N + 1 = 3）时，结论为

$$1/2+3y/2-2y^3$$

当比赛为五局三胜（即 2N + 1 = 5）时，结论为

$$1/2+15y/8-5y^3+6y^5$$

当比赛为七局四胜（即 2N + 1 = 7）时，结论为

$$1/2 + 35y/16 - 35y^3/4 + 21y^5 - 20y^7$$

面对具体问题时，通常我们不要求结论绝对精确。因此，当 y 接近 0 时，我们可以仅仅计算以上公式的前两项。

另外一种数学上常用的分布是“泊松分布”，其名称是根据法国统计学家西蒙·丹尼·泊松的名字命名的。泊松分布的处理对象是发生的频率固定但概率极低的随机事件，比如我们在一片草地里发现一株四片叶子的三叶草，一个加油站的工人在给一辆宝石捷跑车加完油之后感到由衷的激动，19 世纪一名普鲁士军官被服役多年的战马踢死等等。泊松分布其实是二项分布的一个推广，它讨论的是在二项分布中试验次数极多而在每一次试验中 S 出现的概率极低的情况。例如，在一个由报社举办的有奖竞猜活动中，参与者数目极多（但不确定），每个参与者获奖的概率极低，此时获奖者的人数就符合泊松分布，292 获奖者人数与平均获奖人数有关。用 X 表示获奖人数，用 μ 表示平均获奖人数，X 的值可能是 0，1，2，3，……，X = r 的概率为

$$p(r) = \frac{e^{-u}\mu^r}{r!}$$

其中 r = 0，1，2，3，……。表 AII.1 给出了 $\mu = 2.5$ 时的几个概率值。

表 AII.1

r 的值	0	1	2	3	4	5	6	7 或 7 以上
对应的概率值	0.082	0.205	0.257	0.214	0.134	0.067	0.028	0.014

随着 r 的值的增加，最初 p（r）的值递增，逐渐达到最大值，其后逐步递减。这个性质与二项分布相同。

以上我们讨论的对象都是离散量。面对连续量（如时间和距离）时，以上方法失效。例如，彼得的体重为72公斤的概率是多少？这个概率值为0。即使彼得的体重为71.8（甚至精确度更高的71.82）公斤，我们也必须认为他的体重不是72公斤。当讨论对象为连续量时，我们必须换一个分析思路。

正确的思路是，我们不考虑彼得的体重恰好是72公斤的概率（这个概率值只能为0），而是考虑彼得的体重处于某个固定的小区间内的概率，例如，其体重在71.5～72.5公斤内的概率，或者其体重在71.9～72.1公斤内的概率，等等，区间的宽度取决于我们所希望的精确程度，可以任意规定。在处理连续量时，我们只关心一个值处于某一段区域内的概率，而不关心这个值恰好等于某个值的概率。

例如，我们把一组正常的成年男子作为样本，统计其体重的分布情况，并以柱状图表示统计结果，如图AII.1所示。我们把横坐标轴划分为71.5～72.5公斤之类的若干区间，统计每个对象的体重落在各个区间内的频率，最终得到一条光滑曲线，这条曲线就可以表示样本的体重分布规律。这条曲线和横坐标轴围成一个封闭区域，此区域的面积为1。如果我们想知道一个对象的体重落在某个区间内（比如说70～75公斤）的概率，这个概率值可以用一个封闭区域的面积反映，此封闭区域由横坐标轴、光滑曲线和代表规定区间的两条竖线围成，如图AII.2所示。

293

最常用的连续分布是所谓的“正态分布”。正态分布曲线的形状像一只铃铛。图AII.3是一条标准的正态分布曲线，标准的正态分布曲线的特征是关于纵坐标轴对称，而且曲线与横坐标轴围成的面积为1。非标准的正态分布曲线可以转化为标准的正态分布曲线，只要平移纵坐标轴的位置、调整横坐标轴

的单位即可。

图 AⅡ.1：180 个正常成年男子的体重分布（假想数据）

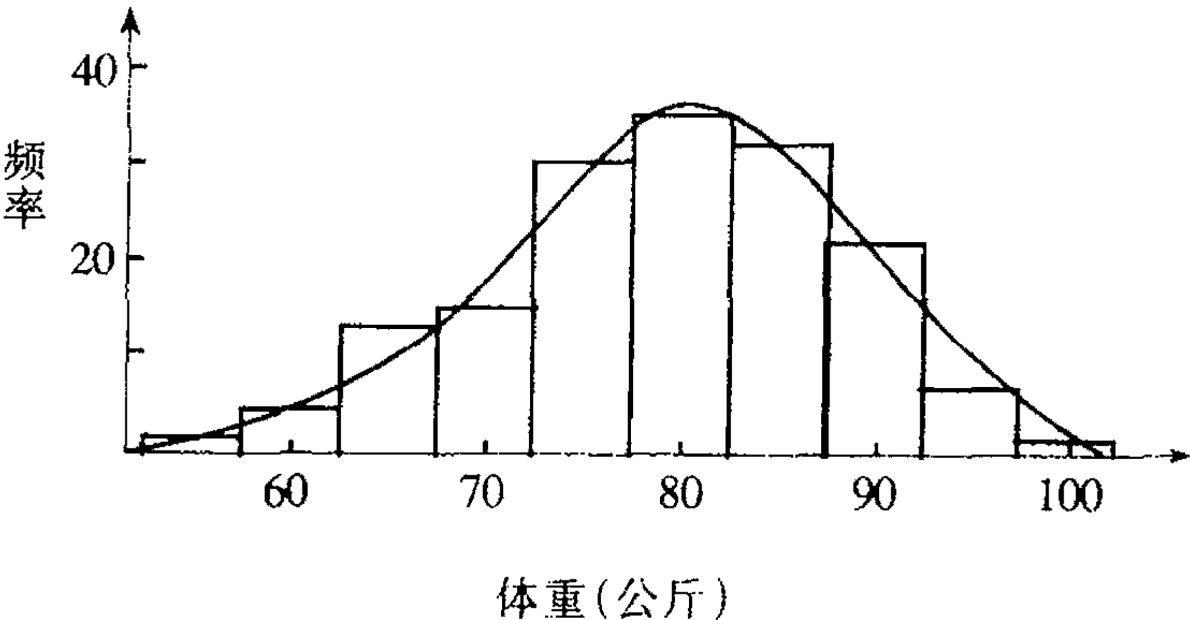
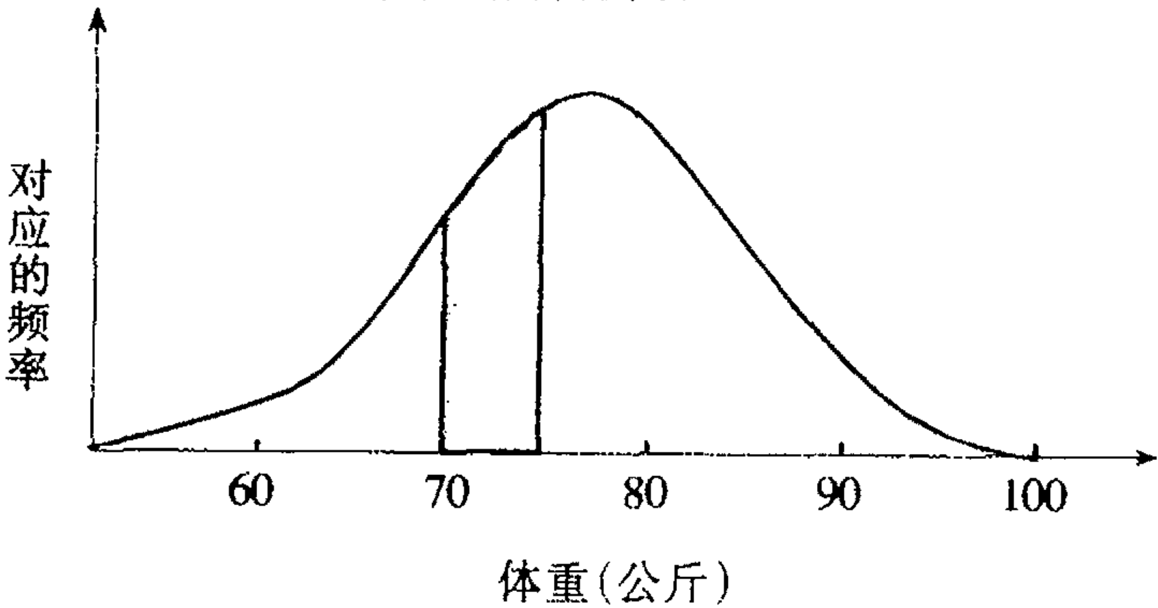


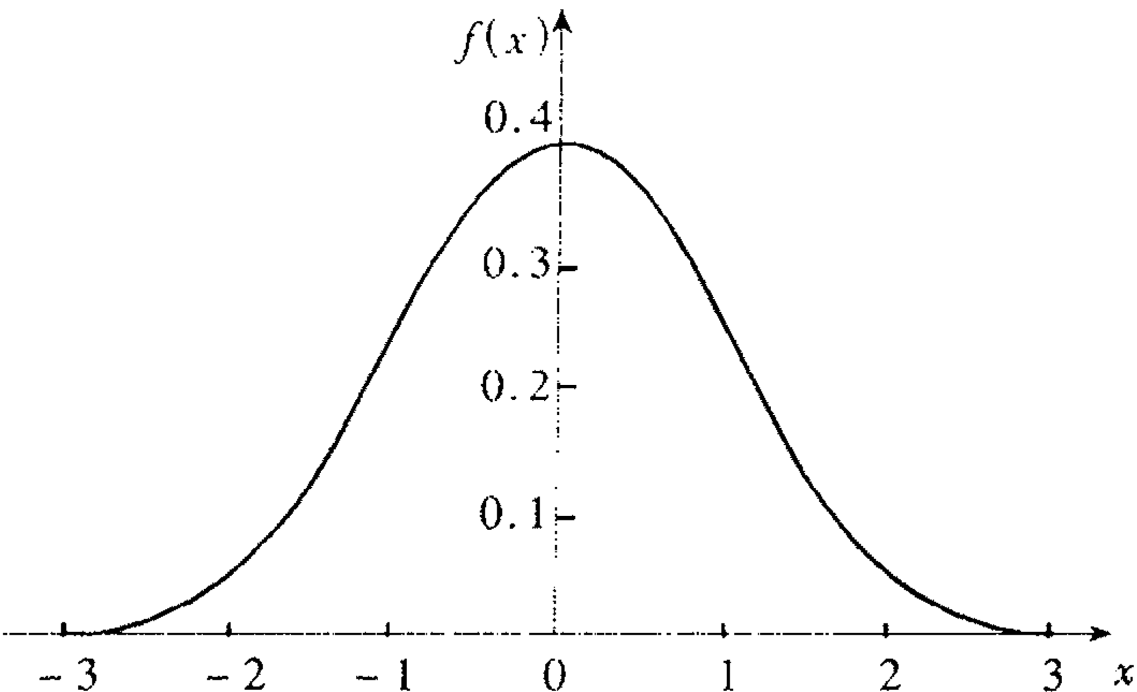
图 AⅡ.2：一个对象的体重位于 70～75 公斤的区间内的概率（阴影部分面积）



数学家详细研究了标准的正态分布曲线的规律，并形成了实用性很强的数学用表。当我们在标准的正态分布曲线下计算某一个概率值时，通常不需要复杂的推演，在数学用表中就可以查到答案。附录 3 中有进一步的介绍。在统计学上正态分布相当重要。

在伯努利试验中，首次 S 出现在第 n 次试验中的概率符合几何分布。应用几何分布的一个例子是分析在一场足球赛中第一个进球发生的时间。假定进球是随机事件，单位时间内进球

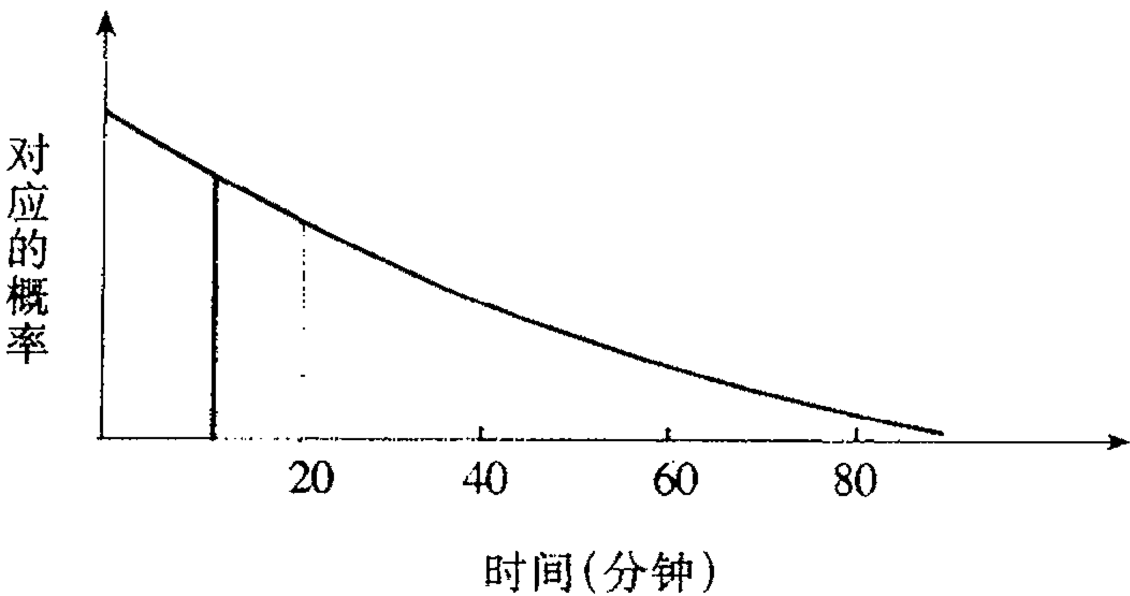
图 A Ⅱ.3：标准正态分布



的概率保持不变，则第一个进球发生的时间可以表示为图 AII.4 的形式，这种情况在数学上称为“指数分布”。在图 AII.4 中，指数分布曲线与横坐标轴围成的面积为 1，这个特征与正态分布曲线相同。如果我们想知道第一个进球发生在第 12 ~ 20 分钟之间的概率，只需要计算一个封闭区域的面积，此封闭区域由横坐标轴、光滑曲线和代表时间区间的两条竖线围成。这种方法可以处理类似问题。与几何分布相同，指数分布是递减的。

295

图 A Ⅱ.4：第一个进球发生在第 12 ~ 20 分钟之间的概率（阴影部分面积）



296

附录3　平均值与变异性

多数人对于“平均值”的理解是直观的。通常，我们使用“平均值”这个词的时候不需要严格的定义。然而，为了研究平均值的应用范围及其局限，我们需要更严格的分析。

概率分布可以告诉我们关于对象的全部信息：有多少种可能结果，每种结果发生的概率是多少。有时候我们确实需要完整的信息，但也有些时候我们发现过分全面的信息反而使我们抓不住重点，我们希望对资料进行有效的综合，提炼出核心因素。这些因素通常包括：

○ 最可能出现的结果；

○ 平均结果；

○ 各种结果关于平均值的变异性。

假设我们参与一个有奖游戏，我们关心的对象是我们在这个游戏中将获得的奖金额 X。我们可以列出一个表，在表中指明各种可能结果及其对应的概率，这个表就是奖金额的概率分布。在游戏中，我们最关心的就是以上三个因素。如果这个游戏是大英帝国赛马会，最可能出现的结果是 $X=0$。

平均值

最简单的情况是 X 只有两个可能值。假设 X = 2 或 X = 8。平均值取决于这两种结果各自对应的概率。假定 X = 2 的概率为 70%，一开始我们的奖金额可能是这样：

2 2 2 8 2 2 8 8 2 2 2 2 8 2 8 2 2 2

显然，平均值接近 2 而远离 8，因为 2 出现的频率较高。如果我们玩了 1 000 次，其中 700次 X = 2 而 300 次 X = 8，则平均值为 [297]

$$\frac{700\times2+300\times8}{1000}=0.7\times2+0.3\times8=3.8$$

在一般情况下，我们就是这样计算平均值的：用每个可能值乘以其对应的概率，而后求和。最后的结果用 E（X）表示，读作“X 的期望值”。我们平常所说的平均值其实就是期望值。

例一：

掷一次骰子，得到的点数的平均值是多少？

解法：

用 X 表示得到的点数，则 X 可能为 1，2，3，4，5，6。每种结果出现的概率为 1/6。于是，

$$E(X)=1\times\frac{1}{6}+2\frac{1}{6}+3\times\frac{1}{6}+4\times\frac{1}{6}+5\times\frac{1}{6}+6\times\frac{1}{6}=3.5$$

当然，实际上 X 的值不可能等于 E（X）。这种情况极其普遍和正常，但有些人对此感到不安，他们认为这种 E（X）

的值缺乏意义。你当然知道这种不安是没道理的，我们可以说“一个家庭平均有 2.4 个孩子”，以及“99%以上的比利时人的腿长的数字高于平均值”，这类说法是有意义的。

平均值具有某些和谐的性质，这些性质通常可以简化计算过程。例如，每个可能值增加一倍，则平均值增加一倍；若干组数据的和的平均值等于每组数据的平均值的和。充分利用这些性质可以使我们节约很多时间。请比较例二的两种解法。

例二：

同时掷两枚骰子，点数之和的平均值是多少？

解法一：

用 X 表示两枚骰子的点数之和，X 的值可能是 2，3，4，……，11，12。每个值对应的概率不同。有三种情况使得

X=4：(1，3)，(2，2) 和 (3，1)

有四种情况使得

X=5：(1，4)，(2，3)，(3，2) 和 (4，1)

有三种情况使得

X=10：(4，6)，(5，5) 和 (6，4)

298 等等。我们可以分别求出每个 X 的值对应的概率，而后求得平均值。如果你在这个繁琐的过程中没犯错误，你的结论将是 7。

解法二：

用 X 表示第一枚骰子的点数，用 Y 表示第二枚骰子的点数，则两枚骰子的点数之和为 X+Y。X 和 Y 的平均值都是 3.5。由于“和的平均值等于平均值的和”，X+Y 的平均值为

3.5 + 3.5 = 7。这个解法简单得多。

例三：

在一局游戏中，你获得的奖金额可能为 1、2 或 3，对应的概率分别为 1/2、1/3、1/6。玩 30 局，平均你将获得多少奖金？

解法：

首先求出在一局游戏中的平均奖金额，然后乘以 30 就得到答案。

在一局游戏中，以 X 表示奖金额，则

$$E(X) = 1 \times \frac{1}{2} + 2 \times \frac{1}{3} + 3 \times \frac{1}{6} = \frac{5}{3}$$

玩 30 局，平均奖金为 30 × 5/3 = 50。

在附录 2 的结论六中，我们介绍了二项分布。计算二项分布中的平均值非常简单。假定一个试验重复了 n 次，在每次试验中结果是 S 的概率为 p，在 n 次试验中平均出现多少次 S?

首先考虑 n = 1 的情况。由于只进行一次试验，可能结果只有两种：出现 1 次 S 或出现 0 次 S，对应的概率分别为 p 和 1 − p。所以 S 出现的平均次数为

$$1 \times p + 0 \times (1 - p) = p$$

由于在 1 次试验中 S 出现的平均次数为 p，所以在 n 次试验中 S 出现的平均次数为 np。

例四：

一个试验重复了 4 次，在每次试验中结果是 S 的概率为 1/3，在 4 次试验中平均出现多少次 S?

299 根据刚刚得出的结论，平均次数为 $4\times1/3=4/3$。

下面我们利用二项分布检验这个结论。根据二项分布公式，S 出现的次数为 0、1、2、3、4 的概率分别为

$$16/81,\ 32/81,\ 24/81,\ 8/81,\ 1/81$$

所以平均次数为

$$0\times16/81+1\times32/81+2\times24/81+3\times8/81+4\times1/81=108/81=4/3$$

可见我们的结论是正确的。

例五及例六：

在第一章中，我们讨论过球队配对问题：16 支球队分 8 对比赛，配对情况尚未公布，你的任务是猜测配对的结果。不关心球队的主客场，只关心你的配对方案与实际是否相符。平均你将猜对多少组配对情况？在第七章中，我们讨论过扑克牌相配问题：把两副扑克牌分别洗乱，把两副牌依次翻开，平均将出现多少次两张牌完全相同的情况？现在我们根据“和的平均值等于平均值的和”这个原则分析这两个问题。

先考虑扑克牌相配问题。首先我们考虑一个特定的位置，比如说两副牌中的第 23 张牌。如果这个位置上的两张牌完全相同，则认为 $X=1$，否则认为 $X=0$。利用这个办法为每个位置分配 X 的值，得到 52 个 X 的值，然后求这些值的和。显然，这个和就是两张牌完全相同的情况发生的次数，于是，问题转化为求这个和的平均值。对于任意一个位置而言，两张牌完全相同的概率为 1/52，所以 X 的平均值为

$$1\times(1/52)+0\times(51/52)=1/51$$

一共有 52 个位置，由于“和的平均值等于平均值的和”，最后的答案为 $52\times(1/52)=1$。即平均将发生 1 次两张牌完

全相同的情况。事实上，无论一副扑克牌包括多少张，答案都是1。

现在考虑球队配对问题。列出球队的实际配对情况，然后与我们的猜测相比较。对于每一个实际的配对，如果我们猜中了这个配对，则认为X=1，否则认为X=0。对于8个实际配对分别确定X的值，然后求这8个值的和。这个和就是我们猜对的配对个数。在我们猜测配对方案的过程中，我们认为每1支球队的对手可能是其他15支球队中的1支，我们随机地选出1个，因此，猜中的概率为1/15。也就是说，对于每个实际配对而言，我们有1/15的机会猜中，X的平均值为1/15。计算8个实际配对对应的平均值的和，答案为8/15。

妥善地利用“和的平均值等于平均值的和”这个原则，很多看似复杂的问题可以迎刃而解。

300

如果球队的总数不是16支，也可以应用以上分析方法，结论是类似的。例如，为24支球队配对，平均将猜中12/23组配对；为4支球队配对，平均将猜中2/3组配对。

在附录2的结论四中，我们讨论过几何分布的一个应用：计算首次S出现在某一次试验中的概率。在实战中，我们通常关心相反的问题。假定在一局游戏中我们获胜的概率为1/10，平均而言为了取得第一次胜利我们需要玩多少局？

例七：

在几何分布中，为了得到第一个S平均需要进行1/p次试验。

证明：

为表述方便，用q表示1－p。假定在第X次试验中S首

次出现，则 X = 1，2，3，4……的概率分别为 p，qp，q^2p，q^3p，……。X 的平均值为

$$E(X) = 1 \times p + 2 \times qp + 3 \times q^2p + 4 \times q^3p + \cdots$$
$$= p(1 + 2q + 3q^2 + 4q^3 + \cdots)$$[①]

对高等数学略有所知的读者都可计算出最后的结果：1/p。

抛一枚硬币，得到正面的概率为 1/2，即 p = 1/2。所以为了得到第一个正面，平均需要抛 2 次。这个结论与我们的直观一致。在大英帝国彩票中，你买一张彩票获得头奖的概率为 1/13 983 816，所以为了中第一次头奖，平均你需买 13 983 816 张彩票。简而言之，如果一件事发生的概率为 p，则平均而言你需要等待 1/p 次。

在附录 2 中我们讨论过泊松分布。在泊松分布中，对象的平均值为 μ。请读者验证。(需要一点高等数学知识。)

以上讨论平均值时，我们对这个词的用法与日常语言相同。有一个概念与平均值密切相关，但意义不同，即所谓的"中位数"。中位数也是一个很常用的量。把一组数据按照大小顺序排列，位于正中的那个值就是中位数。在这组数据中随机地选择一个数，这个数不大于中位数的概率是 1/2，这个数不小于中位数的概率也是 1/2。为了理解中位数与平均值的差别，请考虑这个例子：在一个公司中，有一小批人工资额极高。这些人的工资额对公司雇员的工资额的平均值有影响，但是对公司雇员的工资额的中位数没有影响。

考虑七个数：{1 2 3 4 5 6 49}。中位数为 4。如果把 49 换成 7 或者 700 万，中位数不变，还是 4，但平均值将剧烈变化。最初平均值为 10，把 49 换成 7 以后平均值变成 4，把 49 换成

301

① 原文此处有笔误，qp 和 q^2p 的系数错位——译者注。

700万以后平均值超过100万。

有时候一组数据关于中位数完全对称，此时中位数与平均值相等。如果一组数据关于中位数大致对称，则中位数与平均值接近。

变异性

在大英帝国彩票中，每张彩票对应的奖金额的平均值是45便士，而奖金额的中位数是多少与此无关，购买者用其他方法判断中位数。对于彩票的发行者而言，奖金额的平均值是至关重要的数据，因为平均值决定了发行者的利润。但对于购买者而言，平均值就不那么重要。如果一个购买者每周买一张彩票，则平均每周损失55便士。通常他们并不在乎这点损失。是什么东西吸引他们购买彩票呢？设想改变一下彩票的对奖规则：每张彩票都中奖，而且奖金额固定为44便士，此时每张彩票的奖金额的平均值不变，但是彩票的吸引力完全丧失了。吸引购买者的关键因素在于结果的变异性。

数学家发明了很多描述变异性的方法。最常见的方法是用“方差”描述变异性。方差即每个X的值与平均值的差的平方之平均值。用X_1，X_2，X_3，……，X_n表示X的n个值，其平均值为E（X），则方差为$((X_1-E(X))^2+(X_2-E(X))^2+(X_3-E(X))^2+\cdots\cdots+(X_n-E(X))^2)/n$。以彩票为例，如果各个实际的奖金额与平均值的差较小，则方差较小；302 反之，如果各个实际的奖金额与平均值的差较大，则方差较大。

方差的最小值为0。在彩票中，如果每张彩票的奖金额都

相等，则方差为 0。正如概率值不可能在 0 ~ 1 的范围之外，方差的值不可能小于 0。

对于给定的概率分布很容易求出方差。二项分布的方差为 np（1 - p)，几何分布的方差为（1 - p）/p^2，泊松分布的方差为 μ。

由于方差是平方的平均值，所以方差的单位与 X 的单位不同。在彩票的例子中，X 的单位是英镑或便士，而方差的单位无法确定。为了处理这个问题，有时候我们用“标准差”描述变异性。标准差即方差的平方根。标准差的单位与 X 相同。方差与标准差表达的信息完全相同，前者更容易计算，而后者更容易解释。这两个值经常被用来描述变异性，但是不能传达关于变异性的全部信息。如果概率分布符合以下两个条件，则这两个值非常有效：

条件 A：X 的值关于平均值对称；

条件 B：X 的值与平均值的差越大，则 X 对应的概率值越小。

二项分布、泊松分布和正态分布符合这些条件，而几何分布和指数分布不符合这些条件。

对于正态分布而言，如果知道了平均值和标准差，就可以确定这个分布的全部信息。

用 E 表示平均值，用 δ 表示标准差，则正态分布具备以下三个性质：

○ X 的值出现在（E - δ） ~ （E + δ）的范围内的概率大约为 2/3；

○ X 的值出现在（E - 2δ） ~ （E + 2δ）的范围内的概率大约为 95%；

○ X 的值出现在（E - 3δ） ~ （E + 3δ）的范围内的概率

大约为 99.7%。

这三个性质很常用。请参照附录 2 中的图 AII.3。 303

假设我们进行若干局游戏，我们知道每一局中的奖金额的平均值和总共的奖金额的平均值。关于平均值有一个简洁而实用的性质：和的平均值等于平均值的和。方差具备同样的性质：和的方差等于方差的和。

有时我们会遇到这种情况：游戏进行了很多局，从整体上说，条件 A 和条件 B 是被满足的，而对于每一局游戏而言，这两个条件未必被满足。此时，标准差依然是一个很好的参照值。我们刚才介绍的关于正态分布的三个性质通常可以在具体分析中提供帮助。

例如，在第七章的“搜集卡片”一节中，我们在分析中用到了正态分布的性质。其实我们的研究对象并不符合正态分布，但是这些性质依然成立，原因就在于从整体上说条件 A 和条件 B 是被满足的。

在分析搜集卡片问题时，我们把搜集卡片的这个过程分为若干个子过程，每个子过程的目标是得到一张新卡片。这些子过程是相互独立的，比如说，在搜集第 5 张卡片时我们需要买多少件商品对于在搜集第 11 张卡片时我们需要买多少件商品没有影响。这是一个重要条件。每个子过程符合几何分布。

以一个子过程为例：在我们已经取得第 17 张卡片以后，为了取得第 18 张卡片，平均需要买多少件商品？买 1 件商品取得第 18 张卡片的概率为 $p = 7/24$，所以平均需要买 $1/p = 24/7$ 件商品。这是几何分布的性质。在几何分布中，计算方差的公式为 $(1-p)/p^2$，代入 $p = 7/24$，得到方差的值 $(24 \times 17)/7^2$。利用同样的方法可以分析其他子过程。

一般地，为了搜集到 N 张卡片，平均我们需要买

304

$$N\left(1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{N}\right)$$

件商品。我们需要买的商品数的方差为

$$N^2\left(1+\frac{1}{2^2}+\frac{1}{3^2}+\cdots+\frac{1}{N^2}\right)-N\left(1++\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{N}\right)$$

当N不小于10时，这个值的近似公式为

$$N^2\pi^2/6-N\left(0.577+\ln\left(N\right)\right)$$

这就是计算方差的近似公式。下面考虑标准差。当N不小于20时，$N\left(0.577+\ln\left(N\right)\right)$的值比$N^2\pi^2/6$的值小很多，对这个近似公式开平方时可以忽略减号以后的部分，所以标准差为$N\times\pi/\sqrt{6}$，约等于1.28N。当N=49时，平均值为219，标准差为$1.28\times 49=63$，二者之和为262。这就是大英帝国彩票中的数据。

标准差可以用于分析轮盘赌。假定你玩200局，每局押1枚筹码。此时，你的收益额的概率分布符合条件A和条件B。如果只在数字区下注，由于每1局中你将损失1/37，所以在200局中你的平均收益额为$-200\times(1/36)=-5.4$。收益额的方差取决于你的下注方式。

假定你每次都押某一列中的12个数字（即在第十章中的图10.2的F区下注），则每局中的收益额的方差为2.70，于是200局中的方差为$200\times 2.70=540$，开平方得到标准差23.3。平均值+标准差=18，平均值-标准差=-29。所以，你的收益额在-29至18之间的概率为2/3。收益额高于18的概率和收益额低于-29的概率都是1/6。精确地计算收益额的概率分布非常麻烦，但是利用标准差可以轻松地得到一些有用的结论。

假定你每次都押单个数字，则每局中的收益额的方差大了

很多，为 34.08，于是 200 局中的方差为 200 × 34.08，开平方得到标准差 82.6。平均值 + 标准差 = 77，平均值 - 标准差 = -88。所以，你的收益额在 -29 至 18 之间的概率为 2/3。收益额超出这个范围的概率为 1/3。相比之下，这种下注方式使得结果的不可预测性大大增加。 305

如果每次都押红色数字，则平均收益额为 -2.7，标准差为 14.1。平均值 + 2 × 标准差 = 25，平均值 - 2 × 标准差 = -31，所以，你的收益额在 -31 至 25 之间的概率为 95%。平均值 + 3 × 标准差 = 40，平均值 - 3 × 标准差 = -45，所以，你的收益额在 -45 至 40 之间的概率为 95%。你的损失超过 45 的概率只有 1/740。

如果一个赌徒玩了 N 局，收益额的平均值为 -N/37。用 V 表示每局中的收益额的方差，则在全部 N 局中的方差为 NV。开平方之后，我们发现 N 局的收益额的标准差比每局的收益额的标准差增长了$\sqrt{N}$倍。而 N 局的收益额的平均值比每局的平均值的标准差增长了 N 倍。随着 N 的增加，$\sqrt{N}$增长的速度比 N 增长的速度慢很多，也就是说，随着 N 的增加，结果的可预测性增加。

例如，玩 2 万局，每局下注 1 枚筹码，平均收益额为 -540，标准差为 230。平均值 + 2 × 标准差 = -80，平均值 - 2 × 标准差 = -1 000，所以，你的收益额在 -1 000 至 -80 之间的概率为 95%。在这个过程中，你将赢很多局，但是你在整体上获益的概率只有 1%。

但是在大英帝国彩票、足球彩票和增值彩票中，奖金额的方差非常巨大，除非你买了很多很多张彩票，以上分析方法很难提供有指导意义的信息。绝大多数人穷其一生也不会买这么多彩票。分析这类问题需要更精致的数学工具。

在什么情况下可以应用以上分析手段呢？这个问题没有简单的答案。一般说来，奖金额的对称性越好，奖金额远离平均值的概率越低，以上分析方法的适用性越强。请对比轮盘赌和彩票。这个问题的“变异性”太强了。

306

附录4 拟合良度检验

如何判断一枚硬币出现正面和反面的概率是不是相等的?如何判断一枚骰子是不是公平的?这两个问题考查的对象不同,但解决问题的思路是一样的。基本原则是,在有证据表明作为检验对象的器械(硬币、骰子等等)是有问题的之前,我们假定此器械是没问题的。具体的方法是利用此器械产生一组数据,然后把这组数据与标准器械产生的数据相比较,如果二者符合得较好,则认为检查对象是没问题的。事先我们应当规定好数据的丰富程度。检验过程可分为五步:

(1)假定此器械(硬币,骰子,彩票号码发生器等等)是没问题的;

(2)计算出一个标准器械将产生的平均结果;

(3)比较实际数据与计算出的平均结果的偏差;

(4)计算出一个标准器械的结果具备同样(或更大)偏差的概率p;

(5)如果p很小,则否定(1)的假定,认为此器械是有问题的;否则认为没有足够的证据表明此器械有偏差。

p的值越小,则证据越充分。如果得到一个较大的p的

值，则说明检验结果支持“此器械是没问题的”这一假定。请注意，支持这一假定不等于证明这一假定。实际情况很可能是这样：此器械确实有偏差，但是检验的数据量太小，偏差没有暴露出来。

在数学上，这个过程被称为“拟合良度检验”。以检验一枚硬币为例。抛 100 次硬币，得到 X 次正面和 Y 次反面。显然，Y = 100 - X。我们使用两个字母是为了保证讨论过程的对称性。

步骤（1），假定这枚硬币是公平的，这意味着每抛 1 次硬币正面和反面出现的概率都是 1/2。因此，正面和反面应当各出现 50 次——步骤（2）。

307

步骤（3），根据实际的 X 和 Y 的值计算偏差。用 W 表示偏差，计算公式为

$$W=\frac{(X-50)^2}{50}+\frac{(Y-50)^2}{50}$$

这个值称为“拟合良度统计量”。当 X 和 Y 等于平均值时，W = 0。X 和 Y 距离平均值越远，则 W 的值越大。较小的 W 的值可以作为支持步骤（1）中的假定的证据。

步骤（4），计算一枚标准硬币产生偏差 W 的概率。根据附录 2，把一枚标准硬币抛 100 次，结果符合二项分布，对于给定的 W 的值，得到的偏差为 W 的概率是可以计算的。

假定 X = 55，Y = 45，则 W = 1。一枚标准硬币产生同样（或更大）的偏差的概率为 37%。这个概率值可以认为是足够大的，因为一枚标准硬币有三分之一以上的概率得到同样（或更大）的偏差。W = 1 的检验结果支持最初的假定。

假定 X = 60，Y = 40，则 W = 4。一枚标准硬币产生同样（或更大）的偏差的概率仅为 5.7%。此时，我们有理由怀疑

此硬币是有问题的。假定 $X=65$，$Y=35$，则 $W=9$。一枚标准 308 硬币产生同样（或更大）的偏差的概率仅为1/3%。此时，我们有很好的证据证明此硬币是有问题的。

在这三个假定下，p的值分别为37%、5.7%和1/3%。多大的p的值才算足够的小呢？这个问题没有绝对的标准。我们只能说，p的值越小，我们的怀疑越有根据。以下是根据多年的实际检验总结出的几点指导性原则：

（1）当p大于5%时，我们没有理由怀疑此器械是有问题的；

（2）当p在1%～5%之间时，我们有理由怀疑此器械是有问题的；

（3）当p在1/10%～1%之间时，我们有充分的理由怀疑此器械是有问题的；

（4）当p小于1/10%时，我们有强烈的理由怀疑此器械是有问题的。

需要强调的是，以上原则仅仅是建议性的。在给定p的值和标准器械产生的偏差的平均值的情况下，可以确定对应的W的值。表A4.1是一些常用的数据。

下面考虑比硬币更复杂的器械。掷120次骰子，如何根据实际结果进行判断？步骤（1），假定这枚骰子是公平的。由于每个面出现的概率相等，所以每个面平均应当出现20次——步骤（2）。实际上，六个面出现的次数分别为A，B，C，D，E，F。拟合良度统计量为

$$W=\frac{(A-20)^2}{20}+\frac{(B-20)^2}{20}+\cdots+\frac{(F-20)^2}{20}$$

只有在六个面都出现20次的情况下，W的值为0。实际数据与平均值的偏差越大，则W的值越大。在确定了W的值以

后，下一个任务是计算掷一枚标准骰子120次得到同样（或更大）的偏差的概率p。由于骰子比硬币复杂得多，这个过程很麻烦，幸好我们不必了解计算的细节，因为数学家已经准备好了现成的数学用表，我们可以在表中直接查到答案。一枚标准骰子产生的偏差的平均值为5。如果W的值为11，我们对这枚骰子应当怀有戒心。5%、1%和1/10%对应的W的值分别为11、15和21。W的值越大，我们越有理由怀疑这枚骰子有问题。严格说来，为了使我们的结论可靠，必须把骰子掷很多次。在实际检验中，我们通常采用一个经验标准：在步骤（2）中得到的平均值不小于5。如果这个条件被满足，通常可以认为试验次数已足够多。

在第五章的开头，我们介绍过一个检验骰子的例子。检验进行了四次，每次掷120次。当然，120这个次数太小了，不足以暴露问题。四次检验得到的W的值分别为6.8、4.1、9.8、4.6，其中三个值接近5，而且每个值都不超过11。从这些数据出发，我们可以认为这枚骰子是没问题的。

到目前为止，我们讨论了两种器械：硬币和骰子。在这两个例子中，有一个量非常重要：一个标准器械产生的W的值的平均值，这个值称为“自由度”。我们没有必要研究计算自由度的过程，因为我们可以在数学用表中查到需要的数据。

有时候发现对象的偏差是很困难的。一个关键的因素是我们希望发现的偏差的大小。以硬币为例，假设有一枚有问题的硬币，其正面出现的概率为47%。为了发现这个偏差，需要进行多少次试验？抛100次硬币够不够？

抛100次很难发现问题。对于一枚标准硬币而言，抛100次得到的正面多于60次或少于40次的概率为5%。在100次试验之后，只有当得到的正面的次数多于60或少于40时我们

才开始怀疑硬币有问题。然而，这枚有问题的硬币得到的正面次数的平均值为47，得到正面的次数在40~60的区间内的概率为93%。因此，发现问题的机会很小。

如果进行1 000次试验，结果如何？当得到的正面次数超过532或不足468时，我们开始怀疑这枚硬币。对于这枚硬币而言，这种情况出现的概率为1/2。所以，进行1 000次试验，发现问题的机会为50%。

如果进行1万次试验，结果如何？当得到的正面次数超过5 100或不足4 900时，我们开始怀疑这枚硬币。对于这枚硬币而言，这种情况出现的概率只有1/30 000。所以，进行1万次试验足以发现问题。

310

回顾一下检验过程。首先我们求出W的值，而后确定概率p。W的值越大，则p的值越小，我们越有理由怀疑对象是有问题的。但是这里有一个问题：当p=5%时，一个标准器械有5%的可能产生同样的试验结果，如果我们用这个标准检验100个标准器械，会有5个本来没有问题的器械被我们否定，而且我们不知道是哪5个！

在进行实际检验时，我们有可能犯两个错误：其一，器械原本没问题，我们却认为它是有问题的（错杀）；其二，器械原本有问题，我们却认为它是没问题的（漏网）。在实验次数固定的情况下，如果我们试图减少犯第一个错误的概率，就必然增加犯第二个错误的概率。避免错杀与避免漏网，难以两全。例如，我们本来以p=5%为判断的分界，如果把5%减少到1%，则错杀的概率下降，但漏网的概率上升。为了同时降低两种错误发生的概率，只有一个办法：增加试验次数。

现在研究大英帝国彩票的号码发生器。一个号码发生器每次在49个号码中选出6个，对于一个标准器械而言，每个号

码出现的概率为 6/49——步骤（1）。用 D 表示试验进行的次数，则平均每个号码出现 6D/49 次——步骤（2）。用 A 表示这个平均值，用 X 表示一个号码实际上出现的次数，我们可以计算出 49 个值：

$$\frac{(X-A)^2}{A}$$

W 的值就是这 49 个值的和——步骤（3）。（严格地说，这个 W 的值需要乘以 48/43，因为这个器械在 49 个可能结果中选择 6 个。）

对于一个标准器械而言，W 的值的平均值为 48。概率 5%、1%和 1/10%对应的 W 的值分别为 65、74 和 84。1 次检验至少应当包含多少次试验？根据常用的标准，每个号码出现的平均次数应不低于 5，即至少进行 41 次试验，因为

$$41 \times 6/49 = 5.02 > 5$$

我们把 41 次试验作为 1 次检验，而到目前为止我们已经掌握的大英帝国彩票的数据为 246 轮，把这 246 轮的数据分为 6 次检验，得到的 6 个 W 的值分别为 35.8、38.0、40.2、59.8、44.2 和 42.0。每个 W 的值都低于 65，所以我们有理由认为这个器械是公平的。

为了判断一个器械是否公平，检验其各种可能结果出现的频率是否均匀当然是一个重要的方面，但不是全部。另一个需要考虑的因素是每次产生的结果之间是否独立。如果有一枚硬币，它轮流产生正面和反面，则仅仅通过频率检验我们会认为它非常标准，但它显然不是公平的。

311

为了检验一枚硬币每次产生的结果是否相互独立，我们需要计算一次正面后面紧跟一次反面（或正面）的频率，以及一次反面后面紧跟一次正面（或反面）的频率。例如，我们连续

进行30次试验，得到的结果如下：

正反正反反 反正反反反 正正反反正 正反正反反 正反正反正 正正正反反。

一共出现了14次正面和16次反面。拟合良度统计量为

$$W_1=\frac{(14-15)^2}{15}+\frac{(16-15)^2}{15}=\frac{2}{15}=0.1333$$

显然，这个值是可以接受的。另外，考虑试验中出现的14次正面：其中有5次其后紧跟着1次正面，有9次其后紧跟着1次反面；试验中出现过16次反面：其中有9次其后紧跟着1次正面，有7次其后紧跟着1次反面。（我们认为第30次试验后紧跟着第1次试验。）这些数据反应在下表中：

	其后紧跟的正面	其后紧跟的反面
正面	5	9
反面	9	7

假定这枚硬币是公平的，则1次正面之后紧跟着1次正面、1次正面之后紧跟着1次反面、1次反面之后紧跟着1次正面、1次反面之后紧跟着1次反面这四种情况出现的概率应当相等，平均每种情况出现30/4＝7.5次。用A表示这个平均值。根据上表中的数据可以算出拟合良度统计量

$$W_2=\frac{(5-A)^2}{A}+\frac{(9-A)^2}{A}+\frac{(5-A)^2}{A}+\frac{(7-A)^2}{A}=0.14667$$

由于W_2的值受到在这个检验过程中正面和反面出现的频率的影响，所以W_2的值与W_1的值有关。为了消除这个影响，我们求这两个值的差，得到$W=W_2-W_1=1.33$。这个值可以反应每次试验产生的结果之间是否相互独立。

对于一个标准硬币而言，W 的平均值为 2，概率 5%、1% 和 0.1% 对应的 W 的值分别为 6、9 和 14。我们的检验发现的 W 值为 1.33，所以没有证据显示这枚硬币在每次试验中产生的结果之间存在相互影响。

但是，如果我们的 30 次试验得到这样的结果：正面和反面轮流出现，则结论完全不同。此时 1 次正面之后紧跟着 1 次正面、1 次正面之后紧跟着 1 次反面、1 次反面之后紧跟着 1 次正面、1 次反面之后紧跟着 1 次反面这四种情况出现的次数分别为 0、15、15、0。于是，$W_1 = 0$、$W_2 = 30$、$W = 30$。这个 W 值太大了，我们有理由认为这枚硬币在每次试验中产生的结果之间存在相互影响。

利用同样的方法可以分析骰子。第一步，根据在试验中每个面出现了多少次计算出拟合良度统计量 W_1；第二步，统计一个面之后紧跟着另一个面的情况出现了多少次。由于每次试验可能产生 6 种结果，所以我们需要研究 $6 \times 6 = 36$ 种情况，得到 36 个数据。利用这 36 个数据计算出拟合良度统计量 W_2；最后，根据 $W = W_2 - W_1$ 求出 W 的值。一个标准的骰子产生的 W 值的平均值为 30，概率 5%、1% 和 0.1% 对应的 W 的值分别为 44.51 和 60。为了使检验结论可靠，每次检验至少应当进行 180 次试验。（180 这个值是这样得出的：我们需要考虑 36 种情况，平均每种情况至少应当出现 5 次，所以试验次数不低于 180。）

以上例子的核心在于分析各种可能结果出现的频率是否均匀，如果某些结果出现的频率高于另外一些结果，利用这种方法可以检验出来。这种方法其实有极其广泛的应用，最著名的例子之一就是伟大的遗传学家孟德尔对豌豆的研究。孟德尔的研究持续了很多年，不懈的努力和敏锐的洞察最终导致伟大的

科学发现。孟德尔根据自己的遗传学理论提出一个猜想：把豌豆分为 4 种类型，则 4 种类型出现的概率分别为 9/16、3/16、3/16 和 1/16。为了便于表述，我们假设孟德尔进行了 320 次试验，于是 4 种类型的豌豆应当分别出现 180、60、60 和 20 次。根据试验数据可以计算出拟合良度统计量 W。如果豌豆的遗传行为严格符合孟德尔的理论，W 的平均值应当为 3，概率 5%、1% 和 0.1% 对应的 W 的值分别为 8、11 和 16。试验结论非常理想，孟德尔根据实际统计数据得出的 W 的值接近于 0。（有些科学家认为这个试验结论好得可疑，它相当于把一枚标准硬币抛 1 万次，而恰好得到 5 001 次正面。）40 年之后孟德尔的伟大工作被后人重新发现，从而导致了现代生物学的诞生。

为了应用拟合良度检验方法，我们需要知道三个关键概率——5%、1% 和 0.1%——对应的 W 的值。下表是一些常用数据。 313

表 A4.1

平均值	5%	1%	0.1%	平均值	5%	1%	0.1%
1	3.84	6.64	10.63	10	18.31	23.21	29.59
2	5.99	9.21	13.82	15	25.00	30.58	37.70
3	7.82	11.34	16.27	20	31.41	37.57	45.31
4	9.49	13.28	18.47	30	43.77	50.89	59.70
5	11.07	15.09	20.52	40	55.76	63.69	73.40
8	15.51	20.09	26.12	50	67.50	76.15	86.66

314

附录5　克里策略

在第十一章中，我们研究过一个问题：如果你确信一个游戏对你有利，你应当投入多少赌注？当时我们介绍了由约翰·L·克里提供的最佳策略，现在我们讨论克里策略的数学细节。

赔率为1:1的情况

你最初的资金为F，在1局游戏中你可以任意决定下注额（当然下注额不能超过你现有的资金）。游戏的赔率为1:1，即如果你的下注额为X，则本局游戏结束后要么你输掉数额为X的资金，要么你赢得数额为X的资金。游戏对你有利，以p表示你在1局中获胜的概率，$p > 50\%$。假定你在每局中都按固定的比例在你现有的全部资金中拿出一部分下注，用x表示这个比例。你应当如何选择x的值？

贪婪的赌徒将选择非常大的x值。由于游戏对赌徒有利，赌徒投入的赌注越多，收益就越大。当然，在这种选择下赌徒赢多输少，但是1次失利就会使赌徒元气大伤。一个极端贪婪的赌徒会每次都投入全部资金，在幸运的情况下他的资金额将

稳定地成倍增长，然而，1次失利就会使他血本无归。相反，胆小的赌徒将选择非常小的 x 的值。在这种选择下，赌徒的资金将安全地增长，但增长的速度很缓慢。最佳策略是在这两种选择之间取得均衡。

你每次按 x 的比例下注，如果你获胜，则资金变成原来的 $1+x$ 倍；如果你失败，则资金变成原来的 $1-x$ 倍。在游戏进行 N 局之后，假设你获胜 M 局，失败 M－N 局，则你的资金额变为

$$G=(1+x)^{M}(1-x)^{N-M}F$$

你的资金的增长倍数为 G/F。从长期看你的资金按平均增长率呈指数增长，计算增长倍数的指数，得到

$$\log(G/F)=M\log(1+x)+(N-M)\log(1-x)$$

这个值是 N 局游戏的增长指数，为了得到每1局游戏的平均增长指数，需要把这个值除以 N。由于 $M/N=p$，而 $(N-M)/N=1-p$，每1局游戏的平均增长指数为

$$p\log(1+x)+(1-p)\log(1-x)$$

所谓的最佳策略，就是选择一个适当的 x 值，使得这个平均增长指数取最大值。熟悉数学分析的人可以轻松地找到答案：平均增长指数取最大值的条件是 $x=p-(1-p)$。由于 p 就是你在1局中获胜的概率，而 $1-p$ 就是你在1局中失败的概率，所以 $p-(1-p)$ 就是你在这个游戏中的优势的大小。克里得出的第一个结论是：当赔率为1∶1时，如果游戏对你有利，每次你的下注比例应当等于你在游戏中的优势的大小。

这个策略可以保证你的平均资金增长率最大。如果你选择比克里建议的比例更大或更小的 x 值，将导致平均资金增长率的下降。

例如，当 $p=52\%=0.52$ 时，$1-p=0.48$，$x=4\%$。即每

次投入当前资金额的 4% 为最佳策略。把具体的 p 值和 x 值代入上面的公式，就可以求出平均资金增长率。当 p 略大于 1/2 时，平均资金增长率约等于 $x^2/2$。由于 $x = 4\%$，平均资金增长率为 0.000 8。在 N 次游戏之后，你的资金额变成 1.0008^N。由于在实战中会发生连续赢（或输）若干局的情况，所以实际数据可能有较大波动。

在其他赔率下的情况

假定赔率为 K:1（K 可以大于 1，也可以小于 1），在 1 局中你获胜的概率为 p，游戏对你有利，即 $p(K+1) > 1$。你每次按 x 的比例下注，如果你获胜，则资金变成原来的 $1+Kx$ 倍；如果你失败，则资金变成原来的 $1-x$ 倍。利用同样的计算方法，可以得到在每局中的平均资金增长指数

316

$$p\log(1+Kx) + (1-p)\log(1-x)$$

这个值取最大值的条件是 $x = p - (1-p)/K$。在最佳的下注策略下，平均资金增长率近似等于

$$Kx^2(1-x(K-1))/2$$

例如，如果赔率为 2:1，而每局中你获胜的概率为 40%，则最佳的下注比例为

$$0.4 - (1-0.4)/2 = 0.1 = 10\%$$

即每次你应投入全部资金的 10%。此时，平均资金增长率为 0.008。

赌马策略

在一场赛马中，某匹马获胜的概率为 p，其赔率为 α:1。如果 p×（α+1）>1，则在这匹马上下注对你有利。当然，赌博公司总是精心设计赔率以避免这种情况出现，但有时确实会发生这种情况：有一匹（或更多）马对你有利。此时你应如何下注？

用克里策略可以解决这个问题。假设有5匹马（A，B，C，D，E）参赛，每匹马的胜率分别为40%、30%、20%、8%、2%，赔率分别为13:8、9:4、9:2、5:1、20:1。

为了发现哪匹马对你有利，分别计算出 p（α+1）的值，5个值分别为 0.4×21/8=1.05，0.3×13/4=0.975、0.2×11/2=1.1、0.08×6/1=0.48、0.02×21/1=0.42，按这些值的大小顺序为5匹马排序，顺序为C、A、B、D、E。

假定你的策略是在前k匹马上下注。当确定k以后，我们需要计算两个值：这前k匹马的胜率之和，计为P（k）；以及这前k匹马对应的1/（α+1）的值之和，计为A（k）。k=1 317 就意味着你只在1匹马（C）上下注，P（1）和A（1）分别为20%和2/11；k=2就意味着你只在2匹马（C和A）上下注，P（2）和A（2）分别为20%+40%=60%和2/11+8/21=0.5627；k=3就意味着你只在3匹马（C、A和B）上下注，P（3）和A（3）分别为20%+40%+30%=90%和2/11+8/21+4/13=0.870 46；[①] k=4就意味着你在4匹马（C、A、B

① 原文此处有笔误，8/21误为8/13——译者注。

和 D）上下注，P（4）和 A（4）分别为 98%和 1.037。k>4 的情况不必考虑。一旦我们发现 A（k）的值超过 1，我们就停止这个计算过程。克里策略只涉及那些使得 A（k）的值小于 1 的马。

对于每一个使得 A（k）<1 的 k，分别计算出

$$B(k)=\frac{1-P(k)}{1-A(k)}$$

的值。比较 B（k）与第 k 匹马对应的 p×（α+1）的值，如果 p×（α+1）的值大于 B（k），则在前 k 匹马上下注。找到满足这个条件的最大的 k 值，计为 K。克里策略是在前 K 匹马上下注。

在这个例子中，A（1）<1，B（1）=（1-0.2）/（1-2/11）=0.977，第 1 匹马（C）对应的 p×（+1）的值为 1.1，这个值大于 B（1），所以我们在 C 上下注。B（2）=（1-0.6）/（1-0.56277）=0.91 485，第 2 匹马（A）对应的 p×（α+1）的值为 1.05，这个值大于 B（2），所以我们也在 A 上下注。B（3）=（1-0.9）/（1-0.87 046）=0.772，第 3 匹马（B）对应的 p×（α+1）的值为 0.975，这个值大于 B（3），所以我们也在 B 上下注。请注意，虽然 B 不是对我们有利的马，我们仍然在 B 上下注，这个决定从整体上说对我们有利。由于 A（4）>1，我们不考虑 k>3 的情况。此时，K=3，即我们只在前 3 匹马（C，A 和 B）上下注。

最后一步是确定在每匹马上的下注额。我们在全部资金中拿出 1-B（k）的比例下注，对于第 k 匹马（k=1，2，……，K），下注额的比例为 p（k）-B（K）/（α（k）+1）。在这个例子中，我们拿出全部资金的 0.2-0.772 2/11=6%在 C 上下注，拿出全部资金的 0.4-0.772 8/21=10.6%在 A 上下注，拿

出全部资金的 $0.3 - 0.7724/13 = 6.25\%$ 在 B 上下注。

在第十一章中，我们介绍过这个例子。如果把 B 的赔率改为 2:1，其他条件不变，则最佳策略随之改变。此时我们不应当在 B 上下注。请读者完成分析过程。 318

练习题答案

第四章

1. 得到n次正面和n次反面的概率约等于$1/\sqrt{n\pi}$。

a）抛50次，即$n=25$，概率为$1/\sqrt{25\pi}\approx 11\%$；

b）抛250次，即$n=125$，概率约等于5%；

c）抛1 000次，即$n=500$，概率约等于2.5%。

2. 如果对手选择“正正正正”，你的最佳选择是“反正正正”。除非游戏一开始就出现连续四次正面，否则你必胜，胜率为15/16。如果对手选择了“反反正正”，你应当在“正反反正”和“反反反正”之中选择一种确定为最佳选择。由于前后对称的组合出现的概率较低，所以选择后者更好，即你的最佳选择是“反反反正”。你的胜率为2/3。如果你选择“正反反正”，你的胜率为7/12。

3. 比率依次为1，3，5，7，9。

4. 0.002；0.000 2；0.000 025。

5. 出现0次频率倒转的概率最大。最后平衡点只能出现在抛偶数次硬币之后。由于最后平衡点出现在两端的概率最大，所以答案是0或1 000，二者对应的概率相同。

6. 用p（x）表示恰好出现x次频率倒转的概率，则随着x值的增加，p（x）的值递减。由于b）、c）、d）所包含的p（x）的值的个数都是10，所以无论抛多少次硬币，d）对应的概率高于c）对应的概率，而c）对应的概率高于b）对应的概率。 319

抛N次硬币，收支平衡点的值在$\sqrt{N}/3$附近。[①] 当N＝1 001时，收支平衡点为12，即频率倒转的次数不超过12的概率与超过12的概率大致相等，这说明频率倒转的次数不低于30的概率极低。所以四种情况按概率大小排序，顺序为d）、c）、b）、a）。当N＝10 001时，收支平衡点大于30，这说明频率倒转的次数不低于30的概率大于50%。所以四种情况按概率大小排序，顺序为a）、d）、c）、b）。

第六章

1. 各行中的极小值分别为4、7和3，所以最大极小值为3；各列中的极大值分别为12、10和15，所以最小极大值为10。两个值不同，所以鞍点不存在。如果把“15”换成“6”，则最大极小值和最小极大值都是7，于是（b，C）为鞍点。

2. 在表6.6中，求两列值的差，得到（6，－2）；删去正负号并交换次序，得到（2，6），于是罗伊应当以2/8＝1/4的概率选择正面。类似地，求两行值的差，得到（4，－4）；删

① 原文此处有笔误，$\sqrt{N}/3$误为$\sqrt{N/3}$——译者注。

去正负号并交换次序，得到（4，4），于是克林应当以 1/2 的概率选择正面。

3. 不存在优先选择。如果把“10”换成“9”（或更小的数），则 c 可以删除。

4. 支付矩阵如下：

	B	C
a	4	15
b	9	7

两列值的差为（11，－2），所以甲方应当分别以 2/13 和 11/13 的概率选择 a 和 b；两行值的差为（5，－8），所以乙方应当分别以 8/13 和 5/13 的概率选择 B 和 C。甲方在游戏中的平均收益为 107/13，所以游戏的公平参数是 8.23。

5. 由于一枚骰子只有 6 个面，所以至少需要掷 2 次骰子。先掷 1 次，如果得到的点数为“6”，则不计这次结果，再掷 1 次，直到出现点数不为“6”的情况。假定得到的点数为 A。此后，再掷 1 次，如果得到的点数为“1”、“2”或“3”，则认为答案是 A；否则，认为答案是 A＋5。

6. 支付矩阵如下：

	0	1	2	3	4
0	1	1	1	－1	－1
1	1	1	－1	－1	－1
2	1	－1	－1	－1	1
3	－1	－1	－1	1	1

	0	1	2	3	4
4	-1	-1	1	1	1

对于罗伊而言，0优于1，4优于3，所以1和3这两行可以删除。支付矩阵变为：

	0	1	2	3	4
0	1	1	1	-1	-1
2	1	-1	-1	-1	1
4	-1	-1	1	1	1

对于罗伊而言，1优于0，3优于4，所以0和4这两行可以删除。支付矩阵变为：

	1	2	3
0	1	1	-1
2	-1	-1	-1
4	-1	1	1

现在罗伊又可以删除选择2。克林又可以删除选择2。支付矩阵变为：

	1	3
0	1	-1

	1	3
4	- 1	1

问题还原为表 6.1。这是一个公平游戏。罗伊应当随机地选择 0 或 4，克林应当随机地选择 1 或 3。

7. 克林每次都出 2 枚硬币，每局必胜。如果罗伊同意这个规则，他一定是个傻瓜。

8. 当两只鹰相遇时，每一方赢得资源的概率为 50%，所以每一方取得的资源价值为 6；搏斗使双方各损失 8，所以双方的收益都是 - 2。同理，当两只鸽相遇时，双方的收益都是 6 - 3 = 3。支付矩阵如下：

	鹰	鸽
鹰	- 2	12
鸽	0	3

假定鹰所占的比例为 p，则鹰的平均收益为 - 2p + 12（1 - p） = 12 - 14p，鸽的平均收益为 3 - 3p。当 p = 9/11 时，这两个值相等。所以，按 9/11 的概率选择鹰式策略为进化稳定策略。

9. 用 2x 表示资源的价值，则当两只鹰相遇时，双方的收益都是 x - 8；当两只鸽相遇时，双方的收益都是 x - 3；当鹰与鸽相遇时，鹰的收益是 2x，鸽的收益是 0。在这种情况下，321 如果 x 不小于 8，则鹰式策略总是优于鸽式策略，鸽将消失；如果 x 小于 8，则进化稳定策略由两种策略混合构成，选择鹰

式策略的概率为（x+3）/11。这个结论是合理的：资源的价值越小，选择鹰式策略的价值越低。

第十章

1. 每次都押红色数字，大胆的赌法是这样：如果手中的资金不超过50英镑，则投入全部资金；如果手中的资金超过50英镑，则投入一部分资金，确保如果这一轮获胜则资金恰好达到100英镑。最初资金额为20英镑，每一获胜的概率 $p=73/148$。为了完成目标，前两轮必须获胜，资金达到80英镑，概率为p2。此后可能出现三种情况：（1）再赢1次，完成目标；（2）输1次，紧接着赢1次，完成目标；（3）连续输2次，重回起点。用x表示完成目标的概率，则

$$x=p^2\left(p+(1-p)p+(1-p)^2x\right),$$

于是 $x=0.19\,286$。

第二种赌法：首先用16英镑押六个数字，如果获胜则完成目标，概率为6/37。如果失败（概率为31/37），用剩下的4英镑押数字0，如果获胜则拥有144英镑，其概率为1/37。从整体上说，成功的概率为0.184 8。

一种更好的赌法：首先用16英镑押六个数字。如果获胜，则已完成目标；如果失败，拿出3英镑押数字0。如果第二次下注成功，则完成目标；如果再次失败，还剩下1英镑的赌本。此时，用最后1英镑押1列的3个数字（赔率为1赔11），如果获胜（概率为3/37），则资金变为12英镑。再用12英镑押4个数字（赔率为1赔8），如果获胜（概率为4/37），则完成目标。这样，成功率增加了（3/37）×（4/37）=0.0088。

这种赌法的成功率为 0.1936，比押红色数字更好。

2. 在每 1 轮中，平均你将损失 1/37 英镑。90 轮之后，平均你将损失 2.43 英镑。在第 89 轮结束以后，你赢钱的条件是获胜 5 次（获更多）；在第 91 轮结束以后，你赢钱的条件是获胜 6 次（或更多）。在每 1 轮中，你获胜的概率为 x = 2/37，失败的概率为 y = 35/37。利用附录 4 中介绍的二项分布可以得出结论：

在第 89 轮之后，获胜次数不超过 4 的概率为

$y^{89} + 89y^{88}x + 3916y^{87}x^2 + 113564y^{86}x^3 + 2441626y^{85}x^4$,

这个值约等于 0.47019。所以在第 89 轮结束以后，你赢钱的概率大约是 53%。

在第 91 轮之后，获胜次数不超过 5 的概率为

$y^{91} + 91y^{90}x + 4095y^{89}x^2 + 121485y^{88}x^3 + 2672670y^{87}x^4 +$ 322 $46504458y^{86}x^5$,

这个值约等于 0.630 63。所以在第 89 轮结束以后，你赢钱的概率大约是 37%。

3. 酒鬼安全到家的概率是 20/100 = 20%。平均他将在街上走 2 080 = 1 600 步。把他的步长变成原来的一半，即街长为 200 步，酒馆距离家 40 步。酒鬼安全到家的概率是 40/200 = 20%。平均他将在街上走 40 160 = 6 400 步。

4. 假定你的第一轮赌注为 1 枚筹码。

a) 庄家的牌不是烂牌，但你的牌比他的牌大。你赢得 3 枚筹码；

b) 你输掉 3 枚筹码。你应当增加赌注，但是庄家的牌比你的牌好；

c) 你增加赌注，并且获胜。三带二的赔率为 1 赔 8，你赢 323 得 17 枚筹码。

索　引